



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification <sup>7</sup> : C12N 15/12, C07K 14/47, C12N 5/10, C12P 21/00, C07K 16/18, A61K 38/17, C12Q 1/68</p>	A2	<p>(11) International Publication Number: <b>WO 00/56891</b></p> <p>(43) International Publication Date: 28 September 2000 (28.09.00)</p>														
<p>(21) International Application Number: PCT/US00/07817</p> <p>(22) International Filing Date: 22 March 2000 (22.03.00)</p> <p>(30) Priority Data:</p> <table border="0"> <tr> <td>60/125,537</td> <td>22 March 1999 (22.03.99)</td> <td>US</td> </tr> <tr> <td>60/139,565</td> <td>16 June 1999 (16.06.99)</td> <td>US</td> </tr> </table> <p>(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications</p> <table border="0"> <tr> <td>US</td> <td>60/125,537 (CIP)</td> </tr> <tr> <td>Filed on</td> <td>22 March 1999 (22.03.99)</td> </tr> <tr> <td>US</td> <td>60/139,565 (CIP)</td> </tr> <tr> <td>Filed on</td> <td>16 June 1999 (16.06.99)</td> </tr> </table> <p>(71) Applicant (for all designated States except US): INCYTE PHARMACEUTICALS, INC. [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).</p> <p>(72) Inventors; and</p> <p>(75) Inventors/Applicants (for US only): YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose,</p>		60/125,537	22 March 1999 (22.03.99)	US	60/139,565	16 June 1999 (16.06.99)	US	US	60/125,537 (CIP)	Filed on	22 March 1999 (22.03.99)	US	60/139,565 (CIP)	Filed on	16 June 1999 (16.06.99)	<p>CA 95118 (US). HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive, #12, Mountain View, CA 94040 (US). REDDY, Roopa [IN/US]; 1233 W. McKinley Drive, Sunnyvale, CA 94086 (US). BANDMAN, Olga [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). LU, Dyung, Aina, M. [US/US]; 55 Park Belmont Place, San Jose, CA 95136 (US). AZIMZAI, Yalda [US/US]; 2045 Rock Springs Drive, Hayward, CA 94545 (US). YANG, Junming [CN/US]; 7125 Bark Lane, San Jose, CA 95129 (US).</p> <p>(74) Agents: HAMLET-COX, Diana et al.; Incyte Pharmaceuticals, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).</p> <p>(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p><b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i></p>
60/125,537	22 March 1999 (22.03.99)	US														
60/139,565	16 June 1999 (16.06.99)	US														
US	60/125,537 (CIP)															
Filed on	22 March 1999 (22.03.99)															
US	60/139,565 (CIP)															
Filed on	16 June 1999 (16.06.99)															
<p>(54) Title: HUMAN TRANSMEMBRANE PROTEINS</p> <p>(57) Abstract</p> <p>The invention provides human transmembrane proteins (HTMP) and polynucleotides which identify and encode HTMP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of HTMP.</p>																

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

## HUMAN TRANSMEMBRANE PROTEINS

### TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of human transmembrane  
5 proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cell  
proliferative, immunological, reproductive, smooth muscle, and neurological disorders.

### BACKGROUND OF THE INVENTION

Eukaryotic organisms are distinct from prokaryotes in possessing many intracellular  
10 membrane-bound compartments such as organelles and vesicles. Many of the metabolic reactions  
which distinguish eukaryotic biochemistry from prokaryotic biochemistry take place within these  
compartments. In particular, many cellular functions require very stringent reaction conditions, and  
the organelles and vesicles enable compartmentalization and isolation of reactions which might  
otherwise disrupt cytosolic metabolic processes. The organelles include mitochondria, smooth and  
15 rough endoplasmic reticula, sarcoplasmic reticulum, and the Golgi body. The vesicles include  
phagosomes, lysosomes, endosomes, peroxisomes, and secretory vesicles. Organelles and vesicles  
are bounded by single or double membranes.

Biological membranes surround organelles, vesicles, and the cell itself. Membranes are  
highly selective permeability barriers made up of lipid bilayer sheets composed of phosphoglycerides,  
20 fatty acids, cholesterol, phospholipids, glycolipids, proteoglycans, and proteins. Membranes contain  
ion pumps, ion channels, and specific receptors for external stimuli which transmit biochemical  
signals across the membranes. These membranes also contain second messenger proteins which  
interact with these pumps, channels, and receptors to amplify and regulate transmission of these  
signals.

#### 25 Plasma Membrane Proteins

Plasma membrane proteins (MPs) are divided into two groups based upon methods of protein  
extraction from the membrane. Extrinsic or peripheral membrane proteins can be released using  
extremes of ionic strength or pH, urea, or other disruptors of protein interactions. Intrinsic or integral  
membrane proteins are released only when the lipid bilayer of the membrane is dissolved by  
30 detergent.

The majority of known integral membrane proteins are transmembrane proteins (TM) which  
are characterized by an extracellular, a transmembrane, and an intracellular domain. TM domains are  
typically comprised of 15 to 25 hydrophobic amino acids which are predicted to adopt an  $\alpha$ -helical  
conformation. TM proteins are classified as bitopic (Types I and II) and polytopic (Types III and IV)  
35 (Singer, S.J. (1990) Annu. Rev. Cell Biol. 6:247-96). Bitopic proteins span the membrane once while

polytopic proteins contain multiple membrane-spanning segments. TM proteins carry out a variety of important cellular functions acting as cell-surface receptor proteins involved in signal transduction. These functions are represented by growth and differentiation factor receptors, and receptor-interacting proteins such as Drosophila pecanex and frizzled proteins, LIV-1 protein, NF2 protein, and GNS1/SUR4 eukaryotic integral membrane proteins. TM proteins also act as transporters of ions or metabolites, such as gap junction channels (connexins), and ion channels, and as cell anchoring proteins, such as lectins, integrins, and fibronectins. TM proteins are found in vesicle organelle-forming molecules, such as caveolins; or cell recognition molecules, such as cluster of differentiation (CD) antigens, glycoproteins, and mucins.

Many MPs contain amino acid sequence motifs that serve to localize proteins to specific subcellular sites. Examples of these motifs include PDZ domains, KDEL, RGD, NGR, and GSL sequence motifs, von Willebrand factor A (vWFA) domains, and EGF-like domains. RGD, NGR, and GSL motif-containing peptides have been used as drug delivery agents in targeted cancer treatment of tumor vasculature (Arap, W. et al. (1998) Science, 279:377-380). Furthermore, MPs may also contain amino acid sequence motifs that serve to interact with extracellular or intracellular molecules, such as carbohydrate recognition domains (CRD).

Chemical modification of amino acid residue side chains alters the manner in which MPs interact with other molecules, for example, phospholipid membranes. Examples of such chemical modifications to amino acid residue side chains are covalent bond formation with glycosaminoglycans, oligosaccharides, phospholipids, acetyl and palmitoyl moieties, ADP-ribose, phosphate, and sulphate groups.

RNA encoding membrane proteins may have alternative splice sites which give rise to proteins encoded by the same gene but with different messenger RNA and amino acid sequences. Splice variant membrane proteins may interact with other ligand and protein isoforms.

#### G-Protein Coupled Receptors

G-protein coupled receptors (GPCRs) are a superfamily of integral membrane proteins which transduce extracellular signals. GPCRs include receptors for biogenic amines, lipid mediators of inflammation, peptide hormones, and sensory signal mediators. The structure of these highly-conserved receptors consists of seven hydrophobic transmembrane (serpentine) regions, cysteine disulfide bridges between the second and third extracellular loops, an extracellular N-terminus, and a cytoplasmic C-terminus. Three extracellular loops alternate with three intracellular loops to link the seven transmembrane regions. The most conserved parts of these proteins are the transmembrane regions and the first two cytoplasmic loops. A conserved, acidic-Arg-aromatic residue triplet present in the second cytoplasmic loop may interact with G proteins. A GPCR consensus pattern is characteristic of most proteins belonging to this superfamily

(ExPASy PROSITE document PS00237; and Watson, S. and S. Arkininstall (1994) The G-protein Linked Receptor Facts Book, Academic Press, San Diego, CA, pp 2-6). Mutations and changes in transcriptional activation of GPCR-encoding genes have been associated with neurological disorders such as schizophrenia, Parkinson's disease, Alzheimer's disease, drug addiction, and feeding disorders.

#### Scavenger Receptors

Macrophage scavenger receptors with broad ligand specificity may participate in the binding of low density lipoproteins (LDL) and foreign antigens. Scavenger receptors types I and II are trimeric membrane proteins with each subunit containing a small N-terminal intracellular domain, a transmembrane domain, a large extracellular domain, and a C-terminal cysteine-rich domain. The extracellular domain contains a short spacer domain, an  $\alpha$ -helical coiled-coil domain, and a triple helical collagenous domain. These receptors have been shown to bind a spectrum of ligands, including chemically modified lipoproteins and albumin, polyribonucleotides, polysaccharides, phospholipids, and asbestos (Matsumoto, A. et al. (1990) Proc. Natl. Acad. Sci. USA 87:9133-9137; and Elomaa, O. et al. (1995) Cell 80:603-609). The scavenger receptors are thought to play a key role in atherogenesis by mediating uptake of modified LDL in arterial walls, and in host defense by binding bacterial endotoxins, bacteria, and protozoa.

#### Tetraspan Family Proteins

The transmembrane 4 superfamily (TM4SF) or tetraspan family is a multigene family encoding type III integral membrane proteins (Wright, M.D. and Tomlinson, M.G. (1994) Immunol. Today 15:588). The TM4SF is comprised of membrane proteins which traverse the cell membrane four times. Members of the TM4SF include platelet and endothelial cell membrane proteins, melanoma-associated antigens, leukocyte surface glycoproteins, colonal carcinoma antigens, tumor-associated antigens, and surface proteins of the schistosome parasites (Jankowski, S.A. (1994) Oncogene 9:1205-1211). Members of the TM4SF share about 25-30% amino acid sequence identity with one another. A number of TM4SF members have been implicated in signal transduction, control of cell adhesion, regulation of cell growth and proliferation, including development and oncogenesis, and cell motility, including tumor cell metastasis. Expression of TM4SF proteins is associated with a variety of tumors and the level of expression may be altered when cells are growing or activated.

#### Tumor Antigens

Tumor antigens are surface molecules that are differentially expressed in tumor cells relative to normal cells. Tumor antigens distinguish tumor cells immunologically from normal cells and provide diagnostic and therapeutic targets for human cancers (Takagi, S. et al. (1995) Int. J. Cancer 61: 706-715; Liu, E. et al. (1992) Oncogene 7: 1027-1032).

### Ion Channels

Ion channels are found in the plasma membranes of virtually every cell in the body. For example, chloride channels mediate a variety of cellular functions including regulation of membrane potentials and absorption and secretion of ions across epithelial membranes. When present in intracellular membranes of the Golgi apparatus and endocytic vesicles, chloride channels also regulate organelle pH. (See, e.g., Greger, R. (1988) *Annu. Rev. Physiol.* 50:111-122.) Electrophysiological and pharmacological properties of chloride channels, including ion conductance, current-voltage relationships, and sensitivity to modulators, suggest that different chloride channels exist in muscles, neurons, fibroblasts, epithelial cells, and lymphocytes. Many channels have sites for phosphorylation by one or more protein kinases including protein kinase A, protein kinase C, tyrosine kinase, and casein kinase II, all of which regulate ion channel activity in cells. Inappropriate phosphorylation of proteins in cells has been linked to changes in cell cycle progression and cell differentiation. Changes in the cell cycle have been linked to induction of apoptosis or cancer. Changes in cell differentiation have been linked to diseases and disorders of the reproductive system, immune system, and skeletal muscle.

### Proton Pumps

Proton ATPases are a large class of membrane proteins that use the energy of ATP hydrolysis to generate an electrochemical proton gradient across a membrane. The resultant gradient may be used to transport other ions across the membrane ( $\text{Na}^+$ ,  $\text{K}^+$ , or  $\text{Cl}^-$ ) or to maintain organelle pH. Proton ATPases are further subdivided into the mitochondrial F-ATPases, the plasma membrane ATPases, and the vacuolar ATPases. The vacuolar ATPases establish and maintain an acidic pH within various vesicles involved in the processes of endocytosis and exocytosis (Mellman, I. et al. (1986) *Ann. Rev. Biochem.* 55:663-700). Proton-coupled, 12 membrane-spanning domain transporters such as PEPT 1 and PEPT 2 are responsible for gastrointestinal absorption and for renal reabsorption of peptides using an electrochemical  $\text{H}^+$  gradient as the driving force. Another type of peptide transporter, the TAP transporter, is a heterodimer consisting of TAP 1 and TAP 2 and is associated with antigen processing. Peptide antigens are transported across the membrane of the endoplasmic reticulum by TAP so they can be expressed on the cell surface in association with MHC molecules. Each TAP protein consists of multiple hydrophobic membrane spanning segments and a highly conserved ATP-binding cassette (Boll, M. et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:284-289). Pathogenic microorganisms, such as herpes simplex virus, may encode inhibitors of TAP-mediated peptide transport in order to evade immune surveillance (Marusina, K. and Manaco, J.J. (1996) *Curr. Opin. Hematol.* 3:19-26).

### ABC Transporters

ATP-binding cassette (ABC) transporters, also called the "traffic ATPases", are a superfamily

of membrane proteins that mediate transport and channel functions in prokaryotes and eukaryotes (Higgins, C.F. (1992) *Annu. Rev. Cell Biol.* 8:67-113). ABC proteins share a similar overall structure and significant sequence homology. All ABC proteins contain a conserved domain of approximately two hundred amino acid residues which includes one or more nucleotide binding domains. Mutations in ABC transporter genes are associated with various disorders, such as hyperbilirubinemia II/Dubin-Johnson syndrome, recessive Stargardt's disease, X-linked adrenoleukodystrophy, multidrug resistance, celiac disease, and cystic fibrosis.

#### Membrane Proteins Associated with Intercellular Communication

Intercellular communication is essential for the development and survival of multicellular organisms. Cells communicate with one another through the secretion and uptake of protein signaling molecules. The uptake of proteins into the cell is achieved by endocytosis, in which the interaction of signaling molecules with the plasma membrane surface, often via binding to specific receptors, results in the formation of plasma membrane-derived vesicles that enclose and transport the molecules into the cytosol. The secretion of proteins from the cell is achieved by exocytosis, in which molecules inside of the cell are packaged into membrane-bound transport vesicles derived from the *trans* Golgi network. These vesicles fuse with the plasma membrane and release their contents into the surrounding extracellular space. Endocytosis and exocytosis result in the removal and addition of plasma membrane components, and the recycling of these components is essential to maintain the integrity, identity, and functionality of both the plasma membrane and internal membrane-bound compartments.

Lysosomes are the site of degradation of intracellular material during autophagy and of extracellular molecules following endocytosis. Lysosomal enzymes are packaged into vesicles which bud from the *trans*-Golgi network. These vesicles fuse with endosomes to form the mature lysosome in which hydrolytic digestion of endocytosed material occurs. Lysosomes can fuse with autophagosomes to form a unique compartment in which the degradation of organelles and other intracellular components occurs.

Protein sorting by transport vesicles, such as the endosome, has important consequences for a variety of physiological processes including cell surface growth, the biogenesis of distinct intracellular organelles, endocytosis, and the controlled secretion of hormones and neurotransmitters (Rothman, J.E. and Wieland, F.T. (1996) *Science* 272:227-234). In particular, neurodegenerative disorders and other neuronal pathologies are associated with biochemical flaws during endosomal protein sorting or endosomal biogenesis (Mayer R.J. et al. (1996) *Adv. Exp. Med. Biol.* 389:261-269).

Peroxisomes are organelles independent from the secretory pathway. They are the site of many peroxide-generating oxidative reactions in the cell. Peroxisomes are unique among eukaryotic organelles in that their size, number, and enzyme content vary depending upon organism, cell type,

and metabolic needs (Waterham, H.R. and Cregg, J.M. (1997) *BioEssays* 19:57-66). Genetic defects in peroxisome proteins which result in peroxisomal deficiencies have been linked to a number of human pathologies, including Zellweger syndrome, rhizomelic chondrodysplasia punctata, X-linked adrenoleukodystrophy, acyl-CoA oxidase deficiency, bifunctional enzyme deficiency, classical Refsum's disease, DHAP alkyl transferase deficiency, and acatalasemia (Moser, H.W. and Moser, A.B. (1996) *Ann. NY Acad. Sci.* 804:427-441). In addition, Gartner, J. et al. (1991; *Pediatr. Res.* 29:141-146) found a 22 kDa integral membrane protein associated with lower density peroxisome-like subcellular fractions in patients with Zellweger syndrome.

Normal embryonic development and control of germ cell maturation is modulated by a number of secretory proteins which interact with their respective membrane-bound receptors. Cell fate during embryonic development is determined by members of the activin/TGF- $\beta$  superfamily, cadherins, IGF-2, and other morphogens. In addition, proliferation, maturation, and redifferentiation of germ cell and reproductive tissues are regulated, for example, by IGF-2, inhibins, activins, and follistatins (Petraglia, F. (1997) *Placenta* 18:3-8; Mather, J.P. et al. (1997) *Proc. Soc. Exp. Biol. Med.* 215:209-222).

#### **Endoplasmic Reticulum Membrane Proteins**

The normal functioning of the eukaryotic cell requires that all newly synthesized proteins be correctly folded, modified, and delivered to specific intra- and extracellular sites. Newly synthesized membrane and secretory proteins enter a cellular sorting and distribution network during or immediately after synthesis (cotranslationally or posttranslationally) and are routed to specific locations inside and outside of the cell. The initial compartment in this process is the endoplasmic reticulum (ER) where proteins undergo modifications such as glycosylation, disulfide bond formation, and assembly into oligomers. The modified proteins are then transported through a series of membrane-bound compartments which include the various cisternae of the Golgi complex, where further carbohydrate modifications occur. Transport between compartments occurs by means of vesicles that bud and fuse in a manner specific to the type of protein being transported. Once within the secretory pathway, proteins do not have to cross a membrane to reach the cell surface.

Although the majority of proteins processed through the ER are transported out of the organelle, some are retained. The signal for retention in the ER in mammalian cells consists of the tetrapeptide sequence, KDEL, located at the carboxyl terminus of proteins (Munro, S. (1986) *Cell* 46:291-300). Proteins containing this sequence leave the ER but are quickly retrieved from the early Golgi cisternae and returned to the ER, while proteins lacking this signal continue through the secretory pathway.

Disruptions in the cellular secretory pathway have been implicated in several human diseases. In familial hypercholesterolemia the low density lipoprotein receptors remain in the ER,



rather than moving to the cell surface (Pathak, R.K. (1988) J. Cell Biol. 106:1831-1841). Altered transport and processing of the  $\beta$ -amyloid precursor protein ( $\beta$ APP), involves the putative vesicle transport protein presenilin, and may play a role in early-onset Alzheimer's disease (Levy-Lahad, E. et al. (1995) Science 269:973-977). Changes in ER-derived calcium homeostasis have been

5 associated with diseases such as cardiomyopathy, cardiac hypertrophy, myotonic dystrophy, Brody disease, Smith-McCort dysplasia, and diabetes mellitus.

#### **Mitochondrial Membrane Proteins**

The mitochondrial electron transport (or respiratory) chain is a series of three enzyme complexes in the mitochondrial membrane that is responsible for the transport of electrons from

10 NADH to oxygen and the coupling of this oxidation to the synthesis of ATP (oxidative phosphorylation). ATP then provides the primary source of energy for driving the many energy-requiring reactions of a cell.

Most of the protein components of the mitochondrial respiratory chain are the products of nuclear encoded genes that are imported into the mitochondria, and the remainder are products of

15 mitochondrial genes. Defects and altered expression of enzymes in the respiratory chain are associated with a variety of disease conditions in man, including, for example, neurodegenerative diseases, myopathies, and cancer.

#### **Lymphocyte and Leukocyte Membrane Proteins**

The B-cell response to antigens, which is modulated through receptors, is an essential

20 component of the normal immune system. Mature B cells recognize foreign antigens through B cell receptors (BCR) which are membrane-bound, specific antibodies that bind foreign antigens. The antigen/receptor complex is internalized, and the antigen is proteolytically processed. To generate an efficient response to complex antigens, the BCR, BCR associated proteins, and T cell response are all required. Proteolytic fragments of the antigen are complexed with major histocompatibility

25 complex-II (MHCII) molecules on the surface of the B cells where the complex can be recognized by T cells. In contrast, macrophages and other lymphoid cells present antigens in association with MHCI molecules to T cells. T cells recognize and are activated by the MHCI-antigen complex through interactions with the T cell receptor/CD3 complex, a T cell-surface multimeric protein located in the plasma membrane. T cells activated by antigen presentation secrete a variety of

30 lymphokines that induce B cell maturation and T cell proliferation, and activate macrophages, which kill target cells.

Leukocytes have a fundamental role in the inflammatory and immune response, and include monocytes/macrophages, mast cells, polymorphonucleoleukocytes, natural killer cells, neutrophils, eosinophils, basophils, and myeloid precursors. Leukocyte membrane proteins include members of

35 the CD antigens, N-CAM, I-CAM, human leukocyte antigen (HLA) class I and HLA class II gene

products, immunoglobulins, immunoglobulin receptors, complement, complement receptors, interferons, interferon receptors, interleukin receptors, and chemokine receptors.

Abnormal lymphocyte and leukocyte activity has been associated with acute disorders such as AIDS, immune hypersensitivity, leukemias, leukopenia, systemic lupus, granulomatous disease, and eosinophilia.

#### **Apoptosis-Associated Membrane Proteins**

A variety of ligands, receptors, enzymes, tumor suppressors, viral gene products, pharmacological agents, and inorganic ions have important positive or negative roles in regulating and implementing the apoptotic destruction of a cell. Although some specific components of the apoptotic pathway have been identified and characterized, many interactions between the proteins involved are undefined, leaving major aspects of the pathway unknown.

A requirement for calcium in apoptosis was previously suggested by studies showing the involvement of calcium levels in DNA cleavage and Fas-mediated cell death (Hewish, D.R. and L.A. Burgoyne (1973) *Biochem. Biophys. Res. Comm.* 52:504-510; Vignaux, F. et al. (1995) *J. Exp. Med.* 181:781-786; Oshimi, Y. and S. Miyazaki (1995) *J. Immunol.* 154:599-609). Other studies show that intracellular calcium concentrations increase when apoptosis is triggered in thymocytes by either T cell receptor cross-linking or by glucocorticoids, and cell death can be prevented by blocking this increase (McConkey, D.J. et al. (1989) *J. Immunol.* 143:1801-1806; McConkey, D.J. et al. (1989) *Arch. Biochem. Biophys.* 269:365-370). Therefore, membrane proteins such as calcium channels are important for the apoptotic response.

The discovery of new human transmembrane proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative, immunological, reproductive, smooth muscle, and neurological disorders.

#### **SUMMARY OF THE INVENTION**

The invention features purified polypeptides, human transmembrane proteins, referred to collectively as "HTMP" and individually as "HTMP-1," "HTMP-2," "HTMP-3," "HTMP-4," "HTMP-5," "HTMP-6," "HTMP-7," "HTMP-8," "HTMP-9," "HTMP-10," "HTMP-11," "HTMP-12," "HTMP-13," "HTMP-14," "HTMP-15," "HTMP-16," "HTMP-17," "HTMP-18," "HTMP-19," "HTMP-20," "HTMP-21," "HTMP-22," "HTMP-23," "HTMP-24," "HTMP-25," "HTMP-26," "HTMP-27," "HTMP-28," and "HTMP-29." In one aspect, the invention provides an isolated polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, c) a biologically active

fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-29. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-29.

- 5           The invention further provides an isolated polynucleotide encoding a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, or d) an immunogenic  
10 fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-29. In one alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:30-58.

- Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, b) a naturally occurring amino acid  
15 sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-29. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention  
20 provides a transgenic organism comprising the recombinant polynucleotide.

- The invention also provides a method for producing a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, c) a biologically active fragment of an amino acid sequence  
25 selected from the group consisting of SEQ ID NO:1-29, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-29. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

- 30           Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, or d) an  
35 immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID

NO:1-29.

The invention further provides an isolated polynucleotide comprising a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:30-58, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected  
5 from the group consisting of SEQ ID NO:30-58, c) a polynucleotide sequence complementary to a), or d) a polynucleotide sequence complementary to b). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a) a  
10 polynucleotide sequence selected from the group consisting of SEQ ID NO:30-58, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:30-58, c) a polynucleotide sequence complementary to a), or d) a polynucleotide sequence complementary to b). The method comprises a) hybridizing the sample with a probe comprising at least 16 contiguous nucleotides comprising a  
15 sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 30 contiguous nucleotides. In another alternative, the probe comprises at least 60  
20 contiguous nucleotides.

The invention further provides a pharmaceutical composition comprising an effective amount of a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, c) a biologically active  
25 fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, and a pharmaceutically acceptable excipient. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional HTMP, comprising administering to a patient in need of such treatment the pharmaceutical composition.

30 The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, or  
35 d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ

ID NO:1-29. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a pharmaceutical composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of  
5 treating a disease or condition associated with decreased expression of functional HTMP, comprising administering to a patient in need of such treatment the pharmaceutical composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, b) a naturally occurring amino acid sequence having at least 90%  
10 sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-29, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-29. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the  
15 invention provides a pharmaceutical composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional HTMP, comprising administering to a patient in need of such treatment the pharmaceutical composition.

20 The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:30-58, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

25

#### BRIEF DESCRIPTION OF THE TABLES

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding HTMP.

30 Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods, algorithms, and searchable databases used for analysis of HTMP.

Table 3 shows selected fragments of each nucleic acid sequence; the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or conditions associated with these tissues; and the vector into which each cDNA was  
35 cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding HTMP were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze HTMP, along with applicable descriptions, references, and threshold parameters.

5

## DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing  
10 particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a  
15 reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be  
20 used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

## 25 DEFINITIONS

"HTMP" refers to the amino acid sequences of substantially purified HTMP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of  
30 HTMP. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of HTMP either by directly interacting with HTMP or by acting on components of the biological pathway in which HTMP participates.

An "allelic variant" is an alternative form of the gene encoding HTMP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or  
35 polypeptides whose structure or function may or may not be altered. A gene may have none, one, or

many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

5           “Altered” nucleic acid sequences encoding HTMP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as HTMP or a polypeptide with at least one functional characteristic of HTMP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding HTMP, and improper or unexpected hybridization to allelic variants,  
10 with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding HTMP. The encoded protein may also be “altered,” and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent HTMP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the  
15 residues, as long as the biological or immunological activity of HTMP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine,  
20 isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

          The terms “amino acid” and “amino acid sequence” refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where “amino acid sequence” is recited to refer to an amino acid sequence of a naturally occurring protein molecule, “amino acid sequence” and like terms are not meant to limit the amino  
25 acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

          “Amplification” relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

30           The term “antagonist” refers to a molecule which inhibits or attenuates the biological activity of HTMP. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of HTMP either by directly interacting with HTMP or by acting on components of the biological pathway in which HTMP participates.

35           The term “antibody” refers to intact immunoglobulin molecules as well as to fragments

thereof, such as Fab, F(ab')<sub>2</sub>, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind HTMP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition capable of base-pairing with the "sense" strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic HTMP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" and "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization



between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acid strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding HTMP or fragments of HTMP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using the XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of one or more Incyte Clones and, in some cases, one or more public domain ESTs, using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that, when made, least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

	Original Residue	Conservative Substitution
25	Ala	Gly, Ser
	Arg	His, Lys
	Asn	Asp, Gln, His
	Asp	Asn, Glu
	Cys	Ala, Ser
	Gln	Asn, Glu, His
30	Glu	Asp, Gln, His
	Gly	Ala
	His	Asn, Arg, Gln, Glu
	Ile	Leu, Val
	Leu	Ile, Val
35	Lys	Arg, Gln, Glu
	Met	Leu, Ile
	Phe	His, Met, Leu, Trp, Tyr
	Ser	Cys, Thr
	Thr	Ser, Val
40	Trp	Phe, Tyr
	Tyr	His, Phe, Trp

Val

Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation,  
 5 (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a  
 10 polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide  
 15 from which it was derived.

A "fragment" is a unique portion of HTMP or the polynucleotide encoding HTMP which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment  
 20 used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50% of a polypeptide) as shown in a certain  
 25 defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:30-58 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:30-58, for example, as distinct from any other sequence in the  
 30 same genome. A fragment of SEQ ID NO:30-58 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:30-58 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:30-58 and the region of SEQ ID NO:30-58 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

35 A fragment of SEQ ID NO:1-29 is encoded by a fragment of SEQ ID NO:30-58. A fragment

of SEQ ID NO:1-29 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-29. For example, a fragment of SEQ ID NO:1-29 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-29. The precise length of a fragment of SEQ ID NO:1-29 and the region of SEQ ID NO:1-29 to which the fragment

5 corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a  
10 target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of  
15 reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the  
20 substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

The phrases "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible  
25 way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of  
30 molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent  
35 similarity" between aligned polynucleotide sequence pairs.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at

- 5 <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>.
- 10 The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such default parameters may be, for example:

*Matrix: BLOSUM62*

15 *Reward for match: 1*

*Penalty for mismatch: -2*

*Open Gap: 5 and Extension Gap: 2 penalties*

*Gap x drop-off: 50*

*Expect: 10*

20 *Word Size: 11*

*Filter: on*

- Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at
- 25 least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

- Nucleic acid sequences that do not show a high degree of identity may nevertheless encode
- 30 similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

- The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a
- 35 standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some

alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the hydrophobicity and acidity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default  
5 parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e  
sequence alignment program (described and referenced above). For pairwise alignments of  
polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap  
penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default  
residue weight table. As with polynucleotide alignments, the percent identity is reported by  
10 CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise  
comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.9  
(May-07-1999) with blastp set at default parameters. Such default parameters may be, for example:

*Matrix: BLOSUM62*

15 *Open Gap: 11 and Extension Gap: 1 penalties*

*Gap x drop-off: 50*

*Expect: 10*

*Word Size: 3*

*Filter: on*

20 Percent identity may be measured over the length of an entire defined polypeptide sequence,  
for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for  
example, over the length of a fragment taken from a larger, defined polypeptide sequence, for  
instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least  
150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment  
25 length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be  
used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain  
DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for  
stable mitotic chromosome segregation and maintenance.

30 The term "humanized antibody" refers to antibody molecules in which the amino acid  
sequence in the non-antigen binding regions has been altered so that the antibody more closely  
resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a  
complementary strand through base pairing under defined hybridization conditions. Specific  
35 hybridization is an indication that two nucleic acid sequences share a high degree of identity.

Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Generally, such wash temperatures are selected to be about 5°C to 20°C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. The  $T_m$  is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating  $T_m$  and conditions for nucleic acid hybridization are well known and can be found in Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g.,  $C_0t$  or  $R_0t$  analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide

sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

“Immune response” can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An “immunogenic fragment” is a polypeptide or oligopeptide fragment of HTMP which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term “immunogenic fragment” also includes any polypeptide or oligopeptide fragment of HTMP which is useful in any of the antibody production methods disclosed herein or known in the art.

The term “microarray” refers to an arrangement of distinct polynucleotides on a substrate.

The terms “element” and “array element” in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

The term “modulate” refers to a change in the activity of HTMP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of HTMP.

The phrases “nucleic acid” and “nucleic acid sequence” refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

“Operably linked” refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

“Peptide nucleic acid” (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

“Probe” refers to nucleic acid sequences encoding HTMP, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes.

“Primers” are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target

polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous  
5 nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

10 Methods for preparing and using probes and primers are described in the references, for example Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel et al., 1987, Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis et al., 1990, PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be  
15 derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to  
20 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3  
25 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific  
30 needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and  
35 polynucleotide fragments identified by any of the above selection methods are useful in hybridization



technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be used to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding HTMP, or fragments thereof, or HTMP itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by

different amino acids or nucleotides, respectively.

“Substrate” refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

“Transformation” describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term “transformed” cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A “transgenic organism,” as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, and plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), supra.

A “variant” of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the “BLAST 2 Sequences” tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or greater sequence identity over a certain defined length. A variant may be described as, for example, an “allelic” (as defined above), “splice,” “species,” or “polymorphic” variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of

polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

10 A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence  
15 identity over a certain defined length of one of the polypeptides.

#### THE INVENTION

The invention is based on the discovery of new human transmembrane proteins (HTMP), the polynucleotides encoding HTMP, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative, immunological, reproductive, smooth muscle, and neurological  
20 disorders.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding HTMP. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each HTMP were identified, and column 4 shows the cDNA  
25 libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. The Incyte clones in column 5 were used to assemble the consensus nucleotide sequence of each HTMP and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention:  
30 column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; column 6 shows the identity of each polypeptide and homologous sequences as identified by BLAST analysis; and column 7 shows analytical methods and in some cases, searchable databases to which the  
35 analytical methods were applied. The methods of column 7 were used to characterize each

polypeptide through sequence homology and protein motifs.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding HTMP. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists fragments of the nucleotide sequences of column 1. These  
5 fragments are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:30-58 and to distinguish between SEQ ID NO:30-58 and related polynucleotide sequences. The polypeptides encoded by the specified fragments of SEQ ID NO:30-58 are useful, for example, as immunogenic peptides. Column 3 lists tissue categories which express HTMP as a fraction of total tissues expressing HTMP. Column 4 lists diseases, disorders, or conditions associated with those  
10 tissues expressing HTMP as a fraction of total tissues expressing HTMP. Column 5 lists the vectors used to subclone each cDNA library. Of particular note is the expression of SEQ ID NO:41 in 21 libraries, of which 15 (71%) are associated with nervous tissue.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding HTMP were isolated. Column 1 references the nucleotide SEQ  
15 ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

SEQ ID NO:45 maps to chromosome 3 within the interval from 49.50 to 55.40 centiMorgans. This interval also contains genes associated with disorders of cell proliferation. SEQ ID NO:47 maps to chromosome 7 within the interval from 74.30 to 76.40 centiMorgans. This interval also contains  
20 an EST associated with cell proliferation. SEQ ID NO:50 maps to chromosome 2 within the interval from 111.5 to 115.3 centiMorgans. This interval also contains genes associated with immune response. SEQ ID NO:51 maps to chromosome 11 within the interval from 84.2 to 87.1 centiMorgans. This interval also contains a gene associated with immune response. SEQ ID NO:53 maps to chromosome 13 within the interval from 77.10 to 86.90 centiMorgans. This interval also  
25 contains genes associated with immune response. SEQ ID NO:55 maps to chromosome 1 within the interval from 74.80 to 78.30 centiMorgans. This interval also contains a gene associated with immune response. SEQ ID NO:58 maps to chromosome 15 within the interval from the p-terminus to 25.30 centiMorgans. This interval also contains a gene associated with cell proliferation.

The invention also encompasses HTMP variants. A preferred HTMP variant is one which  
30 has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the HTMP amino acid sequence, and which contains at least one functional or structural characteristic of HTMP.

The invention also encompasses polynucleotides which encode HTMP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected  
35 from the group consisting of SEQ ID NO:30-58, which encodes HTMP. The polynucleotide

sequences of SEQ ID NO:30-58, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding HTMP. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding HTMP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:30-58 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:30-58. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of HTMP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding HTMP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring HTMP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode HTMP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring HTMP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding HTMP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding HTMP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode HTMP and HTMP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding HTMP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:30-58 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.*

5 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler

10 (Perkin-Elmer). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Perkin-Elmer), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995)

15 Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding HTMP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic

25 DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) *PCR Methods Applic.* 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) *Nucleic Acids Res.* 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments

30 adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) *PCR Methods Applic.* 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) *Nucleic Acids Res.*

35 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries

(Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in  
5 length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T)  
10 library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-  
15 specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments  
20 which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode HTMP may be cloned in recombinant DNA molecules that direct expression of HTMP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a  
25 functionally equivalent amino acid sequence may be produced and used to express HTMP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter HTMP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic  
30 oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number  
35 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat.

Biotechnol. 17:259-264; and Crameri, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of HTMP, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then  
5 subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are  
10 optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding HTMP may be synthesized, in whole or in part,  
15 using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, HTMP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the  
20 ABI 431A peptide synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of HTMP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.)  
25 The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY.)

In order to express a biologically active HTMP, the nucleotide sequences encoding HTMP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains  
30 the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding HTMP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding HTMP. Such signals  
35 include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where



sequences encoding HTMP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be

5 provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression  
10 vectors containing sequences encoding HTMP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and  
15 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding HTMP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus);  
20 plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding HTMP. For example, routine cloning,  
25 subcloning, and propagation of polynucleotide sequences encoding HTMP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding HTMP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for  
30 in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of HTMP are needed, e.g. for the production of antibodies, vectors which direct high level expression of HTMP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

35 Yeast expression systems may be used for production of HTMP. A number of vectors

containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 5 1995, supra; Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; and Scorer, C.A. et al. (1994) *Bio/Technology* 12:181-184.)

Plant systems may also be used for expression of HTMP. Transcription of sequences encoding HTMP may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 10 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology 15 (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding HTMP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain 20 infective virus which expresses HTMP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of 25 DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression 30 of HTMP in cell lines is preferred. For example, sequences encoding HTMP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance 35 to a selective agent, and its presence allows growth and recovery of cells which successfully express

the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk<sup>-</sup>* and *apr<sup>-</sup>* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech),  $\beta$  glucuronidase and its substrate  $\beta$ -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding HTMP is inserted within a marker gene sequence, transformed cells containing sequences encoding HTMP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding HTMP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding HTMP and that express HTMP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of HTMP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on HTMP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See,

e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

5           A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding HTMP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding HTMP, or any fragments thereof, may be cloned into a vector  
10   for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for  
15   ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

          Host cells transformed with nucleotide sequences encoding HTMP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence  
20   and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode HTMP may be designed to contain signal sequences which direct secretion of HTMP through a prokaryotic or eukaryotic cell membrane.

          In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of  
25   the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the  
30   American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

          In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding HTMP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric HTMP protein  
35   containing a heterologous moiety that can be recognized by a commercially available antibody may

facilitate the screening of peptide libraries for inhibitors of HTMP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the HTMP encoding sequence and the heterologous protein sequence, so that HTMP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, *supra*, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled HTMP may be achieved *in vitro* using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, <sup>35</sup>S-methionine.

Fragments of HTMP may be produced not only by recombinant means, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, *supra*, pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A peptide synthesizer (Perkin-Elmer). Various fragments of HTMP may be synthesized separately and then combined to produce the full length molecule.

## 25 THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of HTMP and human transmembrane proteins. In addition, the expression of HTMP is closely associated with cell proliferation and the immune response, and with cardiovascular, gastrointestinal, reproductive, and neurological tissues. Therefore, HTMP appears to play a role in cell proliferative, immunological, reproductive, smooth muscle, and neurological disorders. In the treatment of disorders associated with increased HTMP expression or activity, it is desirable to decrease the expression or activity of HTMP. In the treatment of disorders associated with decreased HTMP expression or activity, it is desirable to increase the expression or activity of HTMP.

Therefore, in one embodiment, HTMP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or

activity of HTMP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an immunological disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a reproductive disorder such as a disorder of prolactin production, infertility, including tubal disease, ovulatory defects, and endometriosis, a disruption of the estrous cycle, a disruption of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, an endometrial or ovarian tumor, a uterine fibroid, autoimmune disorders, an ectopic pregnancy, and teratogenesis, cancer of the breast, fibrocystic breast disease, and galactorrhea, a disruption of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia; a smooth muscle disorder such as angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's syndrome, hypertension, hypoglycemia, myocardial infarction, migraine, and pheochromocytoma, and myopathies including cardiomyopathy, encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, and ophthalmoplegia; and a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis

pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal

5 familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system

10 disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia.

15 In another embodiment, a vector capable of expressing HTMP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HTMP including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified HTMP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat

20 or prevent a disorder associated with decreased expression or activity of HTMP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of HTMP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HTMP including, but not limited to, those listed above.

25 In a further embodiment, an antagonist of HTMP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HTMP. Examples of such disorders include, but are not limited to, those cell proliferative, immunological, reproductive, smooth muscle, and neurological disorders described above. In one aspect, an antibody which specifically binds HTMP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism

30 for bringing a pharmaceutical agent to cells or tissues which express HTMP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding HTMP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HTMP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary

35 sequences, or vectors of the invention may be administered in combination with other appropriate

therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of HTMP may be produced using methods which are generally known in the art. In particular, purified HTMP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind HTMP. Antibodies to HTMP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with HTMP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to HTMP have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of HTMP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to HTMP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc.



Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce HTMP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be  
5 generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA  
10 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for HTMP may also be generated. For example, such fragments include, but are not limited to, F(ab)<sub>2</sub> fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab)<sub>2</sub> fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and  
15 easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such  
20 immunoassays typically involve the measurement of complex formation between HTMP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering HTMP epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay  
25 techniques may be used to assess the affinity of antibodies for HTMP. Affinity is expressed as an association constant,  $K_a$ , which is defined as the molar concentration of HTMP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The  $K_a$  determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple HTMP epitopes, represents the average affinity, or avidity, of the antibodies for  
30 HTMP. The  $K_a$  determined for a preparation of monoclonal antibodies, which are monospecific for a particular HTMP epitope, represents a true measure of affinity. High-affinity antibody preparations with  $K_a$  ranging from about  $10^9$  to  $10^{12}$  L/mole are preferred for use in immunoassays in which the HTMP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with  $K_a$  ranging from about  $10^6$  to  $10^7$  L/mole are preferred for use in immunopurification and similar  
35 procedures which ultimately require dissociation of HTMP, preferably in active form, from the

antibody (Catty, D. (1988) Antibodies. Volume I: A Practical Approach, IRL Press, Washington, DC; Liddell, J.E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to  
5 determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of HTMP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g.,  
10 Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding HTMP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding HTMP may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with  
15 sequences complementary to polynucleotides encoding HTMP. Thus, complementary molecules or fragments may be used to modulate HTMP activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding HTMP.

20 Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding HTMP. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

25 Genes encoding HTMP can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding HTMP. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more  
30 with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding HTMP. Oligonucleotides derived from the transcription  
35 initiation site, e.g., between about positions -10 and +10 from the start site, may be employed.

Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in  
5 Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme  
10 molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding HTMP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA,  
15 GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

20 Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding HTMP. Such DNA sequences may be incorporated into a wide variety of vectors  
25 with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3'  
30 ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous  
35 endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

10 An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of HTMP, antibodies to HTMP, and mimetics, agonists, antagonists, or inhibitors of HTMP. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing  
15 compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial,  
20 intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on  
25 techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees,  
30 capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol,  
35 and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose,

hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

5 Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

10 Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or  
15 liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or  
20 dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of  
25 highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating,  
30 dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acids. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preparation may be a lyophilized powder which may contain any  
35 or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a

pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of HTMP, such labeling would include amount, frequency, and method of administration.

5        Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, or pigs.

10       An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example HTMP or fragments thereof, antibodies of HTMP, and agonists, antagonists or inhibitors of HTMP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the  $ED_{50}$  (the dose therapeutically effective in 50% of the population) or  $LD_{50}$  (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the  $LD_{50}/ED_{50}$  ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the  $ED_{50}$  with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

25       The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

35       Normal dosage amounts may vary from about 0.1  $\mu$ g to 100,000  $\mu$ g, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art.

Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

## DIAGNOSTICS

5 In another embodiment, antibodies which specifically bind HTMP may be used for the diagnosis of disorders characterized by expression of HTMP, or in assays to monitor patients being treated with HTMP or agonists, antagonists, or inhibitors of HTMP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for HTMP include methods which utilize the antibody and a label to detect HTMP in human body  
10 fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring HTMP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of HTMP expression. Normal  
15 or standard values for HTMP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibody to HTMP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of HTMP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values.  
20 Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding HTMP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of HTMP may be correlated  
25 with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of HTMP, and to monitor regulation of HTMP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding HTMP or closely related molecules may be used to identify nucleic acid sequences which encode HTMP. The specificity of the probe, whether it is  
30 made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding HTMP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50%  
35 sequence identity to any of the HTMP encoding sequences. The hybridization probes of the subject

invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:30-58 or from genomic sequences including promoters, enhancers, and introns of the HTMP gene.

Means for producing specific hybridization probes for DNAs encoding HTMP include the cloning of polynucleotide sequences encoding HTMP or HTMP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as  $^{32}\text{P}$  or  $^{35}\text{S}$ , or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding HTMP may be used for the diagnosis of disorders associated with expression of HTMP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an immunological disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a reproductive disorder such as a disorder of prolactin production, infertility, including tubal disease, ovulatory defects, and endometriosis, a disruption of the estrous cycle, a disruption of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, an endometrial or ovarian tumor, a uterine fibroid, autoimmune disorders, an ectopic pregnancy, and teratogenesis, cancer of



the breast, fibrocystic breast disease, and galactorrhea, a disruption of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia; a smooth muscle disorder such as angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's syndrome, hypertension, hypoglycemia, myocardial infarction, migraine, and pheochromocytoma, and myopathies including cardiomyopathy, encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, and ophthalmoplegia; and a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia .

The polynucleotide sequences encoding HTMP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered HTMP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding HTMP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding HTMP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding HTMP in the

sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of HTMP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding HTMP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding HTMP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced *in vitro*. Oligomers will preferably contain a fragment of a polynucleotide encoding HTMP, or a fragment of a polynucleotide complementary to the polynucleotide encoding HTMP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

Methods which may also be used to quantify the expression of HTMP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer of interest is

presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray  
5 can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g.,  
10 Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding HTMP may be used  
15 to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price,  
20 C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the  
25 location of the gene encoding HTMP on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as  
30 linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other  
35 gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic

linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among  
5 normal, carrier, or affected individuals.

In another embodiment of the invention, HTMP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes  
10 between HTMP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with HTMP, or fragments thereof,  
15 and washed. Bound HTMP is then detected by methods well known in the art. Purified HTMP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing  
20 antibodies capable of binding HTMP specifically compete with a test compound for binding HTMP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with HTMP.

In additional embodiments, the nucleotide sequences which encode HTMP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on  
25 properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder  
30 of the disclosure in any way whatsoever.

The disclosures of all patents, applications and publications, mentioned above and below, in particular U. S. Ser. No. 60/125,537 and U. S. Ser. No. 60/139,565, are hereby expressly incorporated by reference.

## EXAMPLES

### I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERScript plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), pcDNA2.1 plasmid (Invitrogen, Carlsbad CA), or pINCY plasmid (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 $\alpha$ , DH10B, or ElectroMAX DH10B from Life Technologies.

### II. Isolation of cDNA Clones

Plasmids were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or

without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (LabSystems Oy, Helsinki, Finland).

### III. Sequencing and Analysis

cDNA sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Perkin-Elmer) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Perkin-Elmer) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, *supra*, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VI.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the default parameters specified by the clustal algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and PFAM to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:30-58. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

#### IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, *supra*, ch. 7; Ausubel, 1995, *supra*, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\% \text{ sequence identity} \times \% \text{ maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding HTMP occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

#### 10 V. Chromosomal Mapping of HTMP Encoding Polynucleotides

The cDNA sequences which were used to assemble SEQ ID NO:30-58 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:30-58 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 5). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location.

20 The genetic map locations of SEQ ID NO:45, SEQ ID NO:47, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, and SEQ ID NO:58 are described in The Invention as ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Diseases associated with the public and Incyte sequences located within the indicated intervals are also reported in the Invention where applicable.

#### 30 VI. Extension of HTMP Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:30-58 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30



nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one  
5 extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing  $Mg^{2+}$ ,  $(NH_4)_2SO_4$ , and  $\beta$ -mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme  
10 (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times;  
15 Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100  $\mu$ l PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5  $\mu$ l of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II  
20 (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5  $\mu$ l to 10  $\mu$ l aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates,  
25 digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham  
30 Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase  
35 (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following

parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted  
5 with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

In like manner, the nucleotide sequences of SEQ ID NO:30-58 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an  
10 appropriate genomic library.

#### **VII. Labeling and Use of Individual Hybridization Probes**

Hybridization probes derived from SEQ ID NO:30-58 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide  
15 fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250  $\mu$ Ci of [ $\gamma$ -<sup>32</sup>P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech).  
20 An aliquot containing 10<sup>7</sup> counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16  
25 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

#### **VIII. Microarrays**

A chemical coupling procedure and an ink jet device can be used to synthesize array  
30 elements on the surface of a substrate. (See, e.g., Baldeschweiler, *supra*.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After  
35 hybridization, nonhybridized probes are removed and a scanner used to determine the levels and

patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) *Science* 270:467-470; Shalon, D. et al. (1996) *Genome Res.* 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

#### **IX. Complementary Polynucleotides**

Sequences complementary to the HTMP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring HTMP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of HTMP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the HTMP-encoding transcript.

#### **X. Expression of HTMP**

Expression and purification of HTMP is achieved using bacterial or virus-based expression systems. For expression of HTMP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express HTMP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of HTMP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding HTMP by either homologous recombination or bacterial-mediated

transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, HTMP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from HTMP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified HTMP obtained by these methods can be used directly in the following activity assay.

#### **XI. Demonstration of HTMP Activity**

An assay for HMTP activity measures the expression of HMTP on the cell surface. cDNA encoding HMTP is transfected into an appropriate mammalian cell line. Cell surface proteins are labeled with biotin as described (de la Fuente, M. A. et al. (1997) Blood 90:2398-2405). Immunoprecipitations are performed using HMTP-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting techniques. The ratio of labeled immunoprecipitant to unlabeled immunoprecipitant is proportional to the amount of HMTP expressed on the cell surface.

An alternative assay for HMTP activity is based on a prototypical assay for ligand/receptor-mediated modulation of cell proliferation. This assay measures the amount of newly synthesized DNA in Swiss mouse 3T3 cells expressing HMTP. An appropriate mammalian expression vector containing cDNA encoding HMTP is added to quiescent 3T3 cultured cells using transfection methods well known in the art. The transfected cells are incubated in the presence of [<sup>3</sup>H]thymidine and varying amounts of HMTP ligand. Incorporation of [<sup>3</sup>H]thymidine into acid-precipitable DNA is measured over an appropriate time interval using a tritium radioisotope counter, and the amount incorporated is directly proportional to the amount of newly synthesized DNA. A linear dose-response curve over at least a hundred-fold HMTP ligand concentration range is indicative of receptor

activity. One unit of activity per milliliter is defined as the concentration of HMTP producing a 50% response level, where 100% represents maximal incorporation of [<sup>3</sup>H]thymidine into acid-precipitable DNA (McKay, I. and Leigh, I., eds. (1993) Growth Factors: A Practical Approach, Oxford University Press, New York, NY, p. 73).

- 5           Alternatively, an assay for HTMP activity measures the effect of HMTP expression on the regulation of cell growth. To demonstrate that increased levels of HTMP expression correlates with decreased cell motility and increased cell proliferation, expression vectors encoding HTMP are electroporated into highly motile cell lines, such as U-937 (ATCC CRL 1593), HEL 92.1.7 (ATCC TIB 180) and MAC10, and the motility of the electroporated and control cells are compared.
- 10       Methods for the design and construction of an expression vector capable of expressing HTMP in the desired mammalian cell line(s) chosen are well known to the art. Assays for examining the motility of cells in culture are known to the art (cf Miyake, M. et al. (1991) J. Exp. Med. 174:1347-1354 and Ikeyama, S. et al. (1993) J. Exp. Med. 177:1231-1237). Increasing the level of HTMP in highly motile cell lines by transfection with an HTMP expression vector inhibits or reduces the motility of
- 15       these cell lines, and the amount of this inhibition is proportional to the activity of HTMP in the assay.

## **XII. Functional Assays**

- HTMP function is assessed by expressing the sequences encoding HTMP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice
- 20       include pCMV SPORT plasmid (Life Technologies) and pCR3.1 plasmid (Invitrogen), both of which contain the cytomegalovirus promoter. 5-10  $\mu$ g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2  $\mu$ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish
- 25       transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of
- 30       fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with
- 35       specific antibodies; and alterations in plasma membrane composition as measured by the binding of

fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of HTMP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding HTMP and either CD64 or CD64-GFP.

- 5 CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding HTMP and other genes of interest can be analyzed by  
10 northern analysis or microarray techniques.

### **XIII. Production of HTMP Specific Antibodies**

HTMP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

- 15 Alternatively, the HTMP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

- 20 Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-  
25 HTMP activity by, for example, binding the peptide or HTMP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

### **XIV. Purification of Naturally Occurring HTMP Using Specific Antibodies**

- Naturally occurring or recombinant HTMP is substantially purified by immunoaffinity chromatography using antibodies specific for HTMP. An immunoaffinity column is constructed by  
30 covalently coupling anti-HTMP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

- Media containing HTMP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of HTMP (e.g., high ionic strength  
35 buffers in the presence of detergent). The column is eluted under conditions that disrupt

antibody/HTMP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and HTMP is collected.

**XV. Identification of Molecules Which Interact with HTMP**

HTMP, or biologically active fragments thereof, are labeled with <sup>125</sup>I Bolton-Hunter reagent.

5 (See, e.g., Bolton A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled HTMP, washed, and any wells with labeled HTMP complex are assayed. Data obtained using different concentrations of HTMP are used to calculate values for the number, affinity, and association of HTMP with the candidate molecules.

10 Alternatively, molecules interacting with HTMP are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989, Nature 340:245-246), or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

Various modifications and variations of the described methods and systems of the invention  
15 will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the  
20 scope of the following claims.

Table 1

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	30	1316927	BLADTUT02	452732R1 (TYMNOT02), 1316927H1 (BLADTUT02), 1333144F6 (PANCN07), 1428681F1 (SINTBST01), 2671590F6 (ESOGTUT02)
2	31	1354891	LUNGN09	1354891H1 (LUNGN09), 2847485F6 (DRGLN01), 3290718F6 (BONRFET01)
3	32	1415487	BRAIN012	667705R6 (SCORN01), 1415487H1 (BRAIN012), 2350869H1 (COLSUCT01), 2380229F6 (ISLTN01), 2764913H1 (BRSTN012), 2805580H1 (BLADTUT08), 2921276H1 (SININ04), 2997732H1 (OVRTUT07), 3110058F6 (BRSTTUT15), 3689409F6 (HERAN01), SBAA02719F1, SBAA05351F1, SBAA01910F1, SBAA01637F1
4	33	1693184	COLNNOT23	723435R1 (SYNOOAT01), 1227091T6 (COLNNOT01), 1260955R1 (SYNORAT05) 1801610H1 (COLNNOT27), 3221682H1 (COLNNON03), 3272150H1 (BRAINOT20)
5	34	1695049	COLNNOT23	941876H1 (ADREN03), 1695049F6 and 1695049H1 (COLNNOT23), 1864817F6 (PROSN019), 1867837H1 (SKINBIT01), 2659996F6 (LUNGTUT09), 3561726F6 (SKINNOT05), 3647351H1 (ENDINOT01)
6	35	1802448	COLNNOT27	1510235F1 (LUNGN014), 1562316F1 (SPLNNOT04), 1671093F6 (BMARN03), 1694276F6 (COLNNOT23), 1802448H1 and 1802448T6 (COLNNOT27), 2149391F6 (BRAINOT09), 2705476F6 (PONSAT01), 2804539F6 (BLADTUT08), 4792839H1 (EPIBUNT01), 5397728H1 (LIVRTUT13)
7	36	2057214	BEPINOT01	1271518F6 (TESTTUT02), 1365446R1 and 1365446T1 (SCORN02), 1967159R6 (BRSTN04), 2057214H1 (BEPINOT01), 2173927F6 and 2173927T6 (ENDCN03), 2556249F6 (THYMNOT03), SBDA02536F1
8	37	2448177	THP1NOT03	1718148F6 (UCMCN02), 2448177H1 (THP1NOT03), 2520263H1 (BRAINUT21), SCBA02093V1, SCBA01393V1, SCBA01143V1
9	38	2741701	BRSTTUT14	1648954F6 (PROSTUT09), 3418901H1 (UCMCN04), 2741701F6 (BRSTTUT14), 3247835H1 (SEMVN03), 3869680F6 (BMARNOT03), 1809373F6 (PROSTUT12), 992617R6 (COLNNOT11), SBFA01145F1
10	39	3487228	EPIGN01	3408591F6 (PROSTUS08), 3487228F6 (EPIGN01), 3487228H1 (EPIGN01)
11	40	3671426	KIDNTUT16	3671426H1 and 3671426T6 (KIDNTUT16), SASB01034F1, SASB01530F1, SASB00221F1, SASB01242F1



Table 1 (cont.)

Polyptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
12	41	4020882	BRAXNOT02	2413217R6 (BSTMN02), 4020882H1 (BRAXNOT02), SBXA03463D1, SBXA01344D1, SBXA02823D1, SBXA03612D1, SBXA01807D1
13	42	4069777	KIDNNOT26	4069777F6 and 4069777H1 (KIDNNOT26), 4560189F6 (KERATXT01), 4576433H1 (LUNLTUT02), SBGA05674F1
14	43	5309830	MONOTXT02	1274762F6 (TESTTUT02), 1348268T6 (PROSNOT11), 2110201X13F1 and 2110201X17F1 (BRAITUT03), 5309830H1 (MONOTXT02), SAGA02951F1, SAGA02413F1
15	44	306263	HEARNOT01	04615V1 (HMC1NOT01), 167880R6 (LIVRN01), 306263H1 (HEARNOT01), 306263R6 (HEARNOT01), 307430T6 (HEARNOT01), 1391023F6 (EOSINOT01), 2074431H1 (ISLTNOT01), 4938418H1 (EPIMNON04), 5513730H1 (BRADDIRO1), SCHAO4615V1
16	45	483751	HNT2RAT01	037180H1 (HUVENOB01), 143478R1 (TLYMNOR01), 378814H1 (NEUTFMT01), 483751H1 (HNT2RAT01), 483751R6 (HNT2RAT01), 776822R1 (COLNNOT05), 962802R2 (BRSTTUT03), 1353469F1 (LATRTTUT02)
17	46	727332	SYNOOAT01	727332H1 (SYNOOAT01), 1484843T1 (CORPNOT02), 2349284T6 (COLSUCT01), 2705135H1 (PONSATZ01), 4172701F6 (SINTNOT21), SBYA04629U1, SBYA05701U1, SBYA06282U1, SBYA05724U1
18	47	778800	MYOMNOT01	071616F1 (PLACNOB01), 438858T6 (THYRN01), 778800F1 (MYOMNOT01), 778800H1 (MYOMNOT01), 778935R1 (MYOMNOT01), 2638524F6 (BONTNOT01), 4989972H1 (LIVRTUT11), 4996486H1 (MYEPTXT02), 5609965H1 (MONOTXS05)
19	48	1396995	BRAITUT08	1396995H1 (BRAITUT08), 1396995X318U1 (BRAITUT08), 1400161T6 (BRAITUT08), 3519719R6 (LUNGNON03), 3519719T6 (LUNGNON03)
20	49	1597730	BRAINOT14	1597730H1 (BRAINOT14), 2756450T6 (THPIAZS08), 2898729T6 (THYMNON02)
21	50	1629304	COLNPOT01	898259H1 (BRSTNOT05), 1000577R6 (BRSTNOT03), 1629304F6 (COLNPOT01), 1629304H1 (COLNPOT01), 1807094F6 (SINTNOT13), 2628947H1 (PROSTUT12), 3142908H1 (HNT2AZS07), 3217195H1 (TESTNOT07)

Table 1 (cont.)

Polyptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
22	51	1989585	CORPNOT02	470971T6 (NMLR1DT01), 871981R7 (LUNGAST01), 1602687F6 (BLADNOT03), 1989585H1 (CORPNOT02), 1989585T6 (CORPNOT02), 2250984X312D1 (OVARUT01), 2293025R6 (BRAINON01), 4143919H1 (BRSTTMT01), 4855964H1 (BRSTTUT22)
23	52	2134729	ENDCNOT01	1734329F6 (COLNNOT22), 2134729H1 (ENDCNOT01), 2857991H1 (CONNNOT02), 2884112H1 (SINJNOT02), 2970301F6 (HEAONOT02), 3437438F6 (PENCNOT05), 4061613H1 (BRAINOT21), 4239422H1 (SYNWDIT01), 4896767H1 (LIVRTUT12)
24	53	2299506	BRSTNOT05	1226895T1 (COLNNOT01), 1469923F1 (PANCTUT02), 2299506H1 (BRSTNOT05), 2299506R6 (BRSTNOT05), 2748977F6 (LUNGUTUT11)
25	54	2506558	CONUTUT01	838594H1 (PROSTUT05), 1851144F6 (LUNGFET03), 2506558F6 (CONUTUT01), 2506558H1 (CONUTUT01), 3046030H1 (HEAANOT01), 3475919T6 (LUNGNOT27), 4776339H1 (BRAQNOT01)
26	55	2546025	UTRSNOT11	000451H1 (U937NOT01), 820914T1 (KERANOT02), 1441359F1 (THYRNOT03), 1441359R1 (THYRNOT03), 2546025H1 (UTRSNOT11), 3602668H1 (DRGTNOT01)
27	56	3145660	ENDCNOT04	3069110F6 (UTRSNOR01), 3145660F6 (ENDCNOT04), 3145660H1 (ENDCNOT04)
28	57	4901066	OVARDIT01	1752794F6 (LIVRTUT01), 3221661R6 (COLNNON03), 4901066H1 (OVARDIT01), SBQA03652D1, SBQA04510D1
29	58	5031174	BRSTTMT02	033456F1 (THP1NOB01), 082616R6 (HUVESTB01), 1442589T6 (THYRNOT03), 1506778F6 (BRAITUT07), 1643902F6 (HEARFET01), 1901162H1 (BLADTUT06), 2526581H1 (BRAITUT21), 5031174H1 (BRSTTMT02), 5470948H1 (MCLRUNT01)

Table 2

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Identification/ Homologous Sequences	Analytical Methods and Databases
1	417	T19 T48 S53 T248 T402 T52 T142 S196 T238 T261 T300 T329 S395	N9 N17 N190 N236 N239	V163-L178 (aldo/keto reductase) L374-A392 (transmembrane)	membrane glycoprotein g2114323	MOTIFS BLAST HMM
2	316	S70 S108 T123 S73	N139 N143	M1-P39 (signal peptide)	DP2 (transcription factor) g1698896	MOTIFS BLAST SPSCAN
3	273	S76 T63 T152 S239 S16 T17 S101 S137 S156 T193	N99	I2-C28 (transmembrane) M1-G38 (signal peptide)	Voltage-gated Na <sup>+</sup> channel subtype II g3075512	MOTIFS BLAST HMM SPSCAN
4	222	S9 T48 T215		I62-I79, S155-C177 (transmembrane) M1-A35 (signal peptide)	CAC-1 (beta-casein-like protein) g6482350	MOTIFS HMM SPSCAN BLAST
5	329	T285 T325	N51 N171 N269	Y19-W41, W65-G84, L92-V116, N234-Q257 (transmembrane) K18-I263 (tetraspan) Y19-A42, G60-F115, V237-I263 (tetraspan) C15-A33, L66-F104, A149-G160, N234-I263 (tetraspan) L58-L113 (tetraspan)	Tetraspan protein g3152703	MOTIFS BLAST HMM PFAM PRINTS BLOCKS PROFILES SCAN

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Identification/ Homologous Sequences	Analytical Methods and Databases
6	351	T335 S340 S54 S93 T203 T288 S344	N284	CL61-W183 (transmembrane)	Similar to vanadate resistance protein transmembranous domains g3875131	MOTIFS BLAST HMM
7	489	S291 S10 T51 T52 T135 S185 S244 T416 S143 T224 T252 T279 S379 S483	N105 N121	L263-L284, L399-L420 (Leucine zipper) C23-L41, L65-L82, G111- F134, Y152-W170, L306- L324, F345-M364, W429- F455 (transmembrane)	Transmembrane protein	MOTIFS HMM
8	291	T212 T280 T178 S283	N10	A37-F57, I215-L232, W255- F273 (transmembrane) M1-P55 (signal peptide)	Seven transmembrane protein g1314162	MOTIFS HMM SPSCAN BLAST
9	172	T88		M64-F82 (transmembrane) M1-C42 (signal peptide)	Transmembrane protein	MOTIFS HMM SPSCAN
10	155	T92 S94 S114 S136		V19-W42, L46-E65 (aromatic amino acid permease) G41-F57 (transmembrane)	Transmembrane protein	MOTIFS PRINTS HMM
11	578	T133 S516 S59 S263 T273 T38 T75 T169 T273 S461 S518 Y387		V185-W204, M327-C350 (transmembrane) M1-A18 (signal peptide)	Transmembrane protein g1070391	MOTIFS BLAST HMM SPSCAN

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Identification/ Homologous Sequences	Analytical Methods and Databases
12	313	S65 S101 T114 S137 S165 T167 T258 S289 S50 S64 T207 S222 S223 S254 T277		L182-F201 (transmembrane)	Transmembrane protein	MOTIFS HMM
13	205	S108 S147 T158 T37 T92 T169 S173 T197		M1-S18 (signal peptide) M71-A90 (transmembrane)	Transmembrane protein	MOTIFS SPSCAN HMM
14	371	S78 S165 T245 S354 S24 T290 S297	N110 N269	M1-N21 (signal peptide)	Transmembrane protein	MOTIFS SPSCAN
15	374	S10 S57 T59 T74 S312 S130 T193 S322 Y33	N320	L49-L70 (Leucine zipper gene regulatory pattern) L345-Y367 (transmembrane motif) Q222-L333 (Calponin family repeat motif)	g499184 neuronal protein	MOTIFS HMMER BLAST-DOMO BLIMPS-BLOCKS BLAST-GenBank
16	183	T10 T11 S126 T131 T172 T99 T171	N170	G67-Y89, S104-T127, I138- Y157 (transmembrane motifs)		MOTIFS HMMER
17	190	T14 S20 S70 S59 S102 Y66	N13 N62 N131	S188-I190 (Microbodies C- terminal targeting signal) L169-F187 (transmembrane motif) R31-W154 (transmembrane protein)		MOTIFS HMMER BLAST-PRODOM

Table 2 (cont.)

SEQ ID NO.	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Identification/ Homologous Sequences	Analytical Methods and Databases
18	361	S120 S30 S315	N319	L36-L57 (Leucine zipper gene regulatory pattern) L28-V48 (transmembrane motif)		MOTIFS HMMER
19	97	T25	N20	S63-F80 (transmembrane motif)		MOTIFS HMMER
20	232	S49 S68 T103 S104 T39 T92 S181 S216 Y191	N37	L20-F35 (Crystallins beta and gamma signature) M50-Y76, Y131-L158 (transmembrane motifs)	Protein tyrosine phosphatase-like protein PTPLB g6851256	MOTIFS HMMER BLAST
21	271	S39 T59	N37 N41 N111	M1-T21 (Signal peptide) R178-F196, I218-T244 (transmembrane motifs)	TM6P1 (6 transmembrane domain protein) g6013381	MOTIFS HMMER BLAST
22	267	S16 S54 T83 S198 Y63	N24	L121-Y136, Q146-W166, T177-F195 (transmembrane motifs) P3-Y168, Y119-V258 (transmembrane protein)		MOTIFS HMMER BLAST-PRODOM BLAST-DOMO
23	406	S4 T99 S113 S125 S136 S365 S59 T87 S94 Y281	N31 N112	I288-Y313 (transmembrane motif) P9-N406 (transmembrane protein)	g2358019 Homo sapiens T-cell receptor alpha delta locus	MOTIFS HMMER BLAST-PRODOM BLAST-GenBank

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Identification/ Homologous Sequences	Analytical Methods and Databases
24	318	S56 S254 L315 S66 T172 S193 Y180	N313	F8-G26 (transmembrane motif)		MOTIFS HMMER
25	326	S195 T321 S21 S62 S69 S90 S186 T282 T305 S75 S94 T213 S228 S307 Y135 Y264	N290 N294	L231-I250 (transmembrane motif)		MOTIFS HMMER
26	247	S27 S235 S12 T26 S33 S70 S93 S179		L214-S234 (transmembrane motif)	94128007 hyperion gene	MOTIFS HMMER BLAST-GenBank
27	278	S259		G276-F278 (Microbodies C-terminal targeting signal) F137-L155, L176-G195, I262-F278 (transmembrane motifs)		MOTIFS HMMER
28	320	T280 S285 S86 T96 T131 Y156	N84 N109 N121	L197-L218 (Leucine zipper gene regulatory pattern) M1-A35 (Signal peptide) P23-I45, L208-I227, L254-Y272 (transmembrane motifs)		MOTIFS HMMER SPSCAN
29	360	S43 T138 S143 S312 S351	N244 N323 N349 N356	F10-I29, F151-G170, I280-F299 (transmembrane motifs)		MOTIFS HMMER

Table 3

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
30	148-192	Gastrointestinal (0.280) Reproductive (0.240) Nervous (0.180)	Cell proliferation (0.580) Immune response (0.340)	pINCY
31	217-261	Nervous (0.250) Gastrointestinal (0.200) Reproductive (0.200)	Cell proliferation (0.700) Immune response (0.300)	pINCY
32	597-641	Nervous (0.280) Reproductive (0.220) Gastrointestinal (0.140)	Cell proliferation (0.660) Immune response (0.360)	pINCY
33	219-263	Gastrointestinal (0.354) Reproductive (0.271) Nervous (0.083)	Cell proliferation (0.605) Immune response (0.375)	pINCY
34	923-967	Gastrointestinal (0.182) Nervous (0.182) Cardiovascular (0.159)	Cell proliferation (0.636) Immune response (0.409)	pINCY
35	273-317	Gastrointestinal (0.231) Reproductive (0.212) Musculoskeletal (0.135)	Cell proliferation (0.731) Immune response (0.288)	pINCY
36	435-479	Reproductive (0.222) Gastrointestinal (0.208) Cardiovascular (0.125) Nervous (0.125)	Cell proliferation (0.653) Immune response (0.361)	PSPORT1
37	381-425	Reproductive (0.312) Nervous (0.281) Hematopoietic/Immune (0.188)	Cell proliferation (0.657) Immune response (0.375)	pINCY
38	165-209	Gastrointestinal (0.231) Nervous (0.154) Reproductive (0.154)	Cell proliferation (0.731) Immune response (0.231)	pINCY



Table 3 (cont.)

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
39	137-181	Nervous (0.333) Reproductive (0.333) Cardiovascular (0.333)	Immune response (0.333) Cell proliferation (0.667)	pINCY
40	456-500	Reproductive (0.324) Nervous (0.206) Hematopoietic/Immune (0.118)	Cell proliferation (0.706) Immune response (0.353)	pINCY
41	354-398	Nervous (0.714) Developmental (0.095) Dermatologic (0.048)	Cell proliferation (0.619) Immune response (0.381)	pINCY
42	273-317	Urologic (0.333) Cardiovascular (0.167) Dermatologic (0.167) Nervous (0.167) Reproductive (0.167)	Cell proliferation (0.834) Neurological (0.167)	pINCY
43	149-193	Reproductive (0.417) Urologic (0.167) Cardiovascular (0.083)	Cell proliferation (0.916) Immune response (0.334)	pINCY
44	543-587	Reproductive (0.192) Gastrointestinal (0.154) Nervous (0.154)	Cancer (0.385) Cell proliferation (0.308) Inflammation (0.077)	PBLUESCRIPT
45	545-589	Reproductive (0.220) Cardiovascular (0.171) Hematopoietic/Immune (0.159)	Cancer (0.366) Cell proliferation (0.354) Inflammation (0.293)	PBLUESCRIPT
46	165-209	Nervous (0.486) Gastrointestinal (0.143) Reproductive (0.143)	Cancer (0.343) Inflammation (0.228) Trauma (0.200)	PSPORT1

Table 3 (cont.)

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
47	327-371	Cardiovascular (0.213) Reproductive (0.213) Gastrointestinal (0.128) Nervous (0.128)	Cancer (0.532) Cell proliferation (0.170) Inflammation (0.106)	PSPORT1
48	541-585	Cardiovascular (0.400) Hematopoietic/Immune (0.200) Nervous (0.200) Reproductive (0.200)	Cancer (0.800)	pINCY
49	326-370	Reproductive (0.278) Hematopoietic/Immune (0.167) Gastrointestinal (0.167)	Cancer (0.500) Inflammation (0.222) Cell proliferation (0.167)	pINCY
50	304-351	Reproductive (0.378) Gastrointestinal (0.178) Hematopoietic/Immune (0.089) Urologic (0.089)	Cancer (0.489) Inflammation (0.222) Cell proliferation (0.178)	pINCY
51	36-80	Nervous (0.273) Cardiovascular (0.182) Reproductive (0.227)	Cancer (0.591) Inflammation (0.182) Cell proliferation (0.045)	pINCY
52	236-280	Gastrointestinal (0.273) Nervous (0.212) Reproductive (0.182)	Cancer (0.364) Inflammation (0.242) Cell proliferation (0.121)	pINCY
53	435-479	Reproductive (0.303) Cardiovascular (0.167) Nervous (0.152)	Cancer (0.500) Inflammation (0.152) Trauma (0.106)	PSPORT1
54	326-370	Reproductive (0.276) Hematopoietic/Immune (0.207) Cardiovascular (0.143)	Cancer (0.379) Inflammation (0.276) Cell proliferation (0.138)	pINCY

Table 3 (cont.)

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
55	834-878	Reproductive (0.261) Nervous (0.188) Cardiovascular (0.159)	Cancer (0.333) Cell proliferation (0.218) Inflammation (0.188)	pINCY
56	273-317	Cardiovascular (0.417) Nervous (0.250) Reproductive (0.167)	Inflammation (0.333) Cancer (0.250) Cell proliferation (0.167)	pINCY
57	327-371	Gastrointestinal (0.400) Reproductive (0.200)	Cancer (0.560) Inflammation (0.240) Cell proliferation (0.120)	pINCY
58	444-488	Nervous (0.180) Hematopoietic/Immune (0.156) Reproductive (0.156)	Cancer (0.426) Inflammation (0.221) Cell proliferation (0.131)	pINCY

Table 4

Polynucleotide SEQ ID NO:	Library	Library Comment
30	BLADTUT02	Library was constructed using RNA isolated from bladder tumor tissue removed from an 80-year-old Caucasian female during a radical cystectomy and lymph node excision. Pathology indicated grade 3 invasive transitional cell carcinoma. Family history included acute renal failure, osteoarthritis, and atherosclerosis.
31	LUNGNOT09	Library was constructed using RNA isolated from the lung tissue of a 23-week-old Caucasian male fetus. The pregnancy was terminated following a diagnosis by ultrasound of infantile polycystic kidney disease.
32	BRAINOT12	Library was constructed using RNA isolated from brain tissue removed from the right frontal lobe of a 5-year-old Caucasian male during a hemispherectomy. Pathology indicated extensive polymicrogyria and mild to moderate gliosis (predominantly subpial and subcortical), which are consistent with chronic seizure disorder. Family history included a cervical neoplasm.
33	COLNNOT23	Library was constructed using RNA isolated from diseased colon tissue removed from a 16-year-old Caucasian male during a total colectomy with abdominal/perineal resection. Pathology indicated gastritis and pancolitis consistent with the acute phase of ulcerative colitis. Inflammation was more severe in the transverse colon, with inflammation confined to the mucosa. There was only mild involvement of the ascending and sigmoid colon, and no significant involvement of the cecum, rectum, or terminal ileum. Family history included irritable bowel syndrome.
34	COLNNOT23	Library was constructed using RNA isolated from diseased colon tissue removed from a 16-year-old Caucasian male during a total colectomy with abdominal/perineal resection. Pathology indicated gastritis and pancolitis consistent with the acute phase of ulcerative colitis. Inflammation was more severe in the transverse colon, with inflammation confined to the mucosa. There was only mild involvement of the ascending and sigmoid colon, and no significant involvement of the cecum, rectum, or terminal ileum. Family history included irritable bowel syndrome.

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Library	Library Comment
35	COLNNOT27	Library was constructed using RNA isolated from diseased cecal tissue removed from 31-year-old Caucasian male during a total intra-abdominal colectomy, appendectomy, and permanent ileostomy. Pathology indicated severe active Crohn's disease involving the colon from the cecum to the rectum. There were deep rake-like ulcerations which spared the intervening mucosa. The ulcers extended into the muscularis, and there was transmural inflammation. Patient history included an irritable colon. Previous surgeries included a colonoscopy.
36	BEPINOT01	Library was constructed using RNA isolated from a bronchial epithelium primary cell line derived from a 54-year-old Caucasian male.
37	THP1NOT03	Library was constructed using RNA isolated from untreated THP-1 cells. THP-1 (ATCC TIB 202) is a human promonocyte line derived from the peripheral blood of a 1-year-old Caucasian male with acute monocytic leukemia.
38	BRSTTUT14	Library was constructed using RNA isolated from breast tumor tissue removed from a 62-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated an invasive grade 3 (of 4), nuclear grade 3 (of 3) adenocarcinoma, ductal type. Ductal carcinoma in situ, comedo type, comprised 60% of the tumor mass. Metastatic adenocarcinoma was identified in one (of 14) axillary lymph nodes with no perinodal extension. Tumor cells were strongly positive for estrogen receptors and weakly positive for progesterone receptors. Patient history included benign colon neoplasm, hyperlipidemia, cardiac dysrhythmia, and obesity. Family history included atherosclerotic coronary artery disease, myocardial infarction, colon cancer, ovarian cancer, lung cancer, and cerebrovascular disease.
39	EPIGNOT01	Library was constructed using RNA isolated from epiglottic tissue removed from a 71-year-old male during laryngectomy with right parathyroid biopsy. Pathology for the associated tumor tissue indicated recurrent grade 1 papillary thyroid carcinoma.

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Library	Library Comment
40	KIDNTUT16	Library was constructed using RNA isolated from left pole kidney tumor tissue removed from a 53-year-old Caucasian female during a nephroureterectomy. Pathology indicated grade 2 renal cell carcinoma. Patient history included hyperlipidemia, cardiac dysrhythmia, metrorrhagia, cerebrovascular disease, atherosclerotic coronary artery disease. Family history included cerebrovascular disease and atherosclerotic coronary artery disease.
41	BRAXNOT02	Library was constructed using RNA isolated from cerebellar tissue removed from a 64-year-old male. Patient history included carcinoma of the left bronchus.
42	KIDNNOT26	Library was constructed using RNA isolated from left kidney medulla and cortex tissue removed from a 53-year-old Caucasian female during a nephroureterectomy. Pathology for the associated tumor tissue indicated grade 2 renal cell carcinoma involving the lower pole of the kidney. Patient history included hyperlipidemia, cardiac dysrhythmia, menorrhagia, normal delivery, cerebrovascular disease, and atherosclerotic coronary artery disease. Family history included cerebrovascular disease and atherosclerotic coronary artery disease.
43	MONOTXT02	Library was constructed using RNA isolated from treated monocytes from peripheral blood removed from a 42-year-old female. The cells were treated with interleukin-10 (IL-10) and lipopolysaccharide (LPS). IL-10 was added at time 0 at 10 ng/ml, LPS was added at 1 hour at 5 ng/ml. The monocytes were isolated from buffy coat by adherence to plastic. Incubation time was 24 hours.
44	HEARNOT01	Library was constructed using RNA isolated from the whole heart tissue of a 56-year-old male, who died from an intracranial bleed.
45	HNT2RAT01	Library was constructed at Stratagene (STR937231), using RNA isolated from the hNT2 cell line (derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor). Cells were treated with retinoic acid for 24 hours

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Library	Library Comment
46	SYNOOAT01	Library was constructed using RNA isolated from the knee synovial membrane tissue of an 82-year-old female with osteoarthritis.
47	MYOMNOT01	Library was constructed using RNA isolated from uterine myometrial tissue removed from a 43-year-old Caucasian female during a vaginal hysterectomy and removal of the fallopian tubes and ovaries. Family history included lung cancer, stroke, type II diabetes, hepatic lesion, chronic liver disease, hyperlipidemia, congenital heart anomaly, and mitral valve prolapse.
48	BRAITUT08	Library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe of a 47-year-old Caucasian male during excision of cerebral meningeal tissue. Pathology indicated grade 4 fibrillary astrocytoma with focal tumoral radionecrosis. Patient history included cerebrovascular disease, deficiency anemia, hyperlipidemia, epilepsy, and tobacco use. Family history included cerebrovascular disease and a malignant prostate neoplasm.
49	BRAINOT14	Library was constructed using RNA isolated from brain tissue removed from the left frontal lobe of a 40-year-old Caucasian female during excision of a cerebral meningeal lesion. Pathology for the associated tumor tissue indicated grade 4 gemistocytic astrocytoma.
50	COLNPOT01	Library was constructed using RNA isolated from colon polyp tissue removed from a 40-year-old Caucasian female during a total colectomy. Pathology indicated an inflammatory pseudopolyp; this tissue was associated with a focally invasive grade 2 adenocarcinoma and multiple tubovillous adenomas. Patient history included a benign neoplasm of the bowel.
51	CORPNOT02	Library was constructed using RNA isolated from diseased corpus callosum tissue removed from the brain of a 74-year-old Caucasian male who died from Alzheimer's disease.
52	ENDCNOT01	Library was constructed using RNA isolated from endothelial cells removed from the coronary artery of a 58-year-old Hispanic male.

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Library	Library Comment
53	BRSTNOT05	Library was constructed using RNA isolated from breast tissue removed from a 58-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology for the associated tumor tissue indicated multicentric invasive grade 4 lobular carcinoma. Patient history included skin cancer, rheumatic heart disease, osteoarthritis, and tuberculosis. Family history included cerebrovascular and cardiovascular disease, breast and prostate cancer, and type I diabetes.
54	CONUTUT01	Library was constructed using RNA isolated from sigmoid mesentery tumor tissue obtained from a 61-year-old female during a total abdominal hysterectomy and bilateral salpingo-oophorectomy with regional lymph node excision. Pathology indicated a metastatic grade 4 malignant mixed mullerian tumor present in the sigmoid mesentery at two sites.
55	UTRSNOT11	Library was constructed using RNA isolated from uterine myometrial tissue removed from a 43-year-old female during a vaginal hysterectomy and removal of the fallopian tubes and ovaries. Pathology for the associated tumor tissue indicated that the myometrium contained an intramural and a submucosal leiomyoma. Family history included benign hypertension, hyperlipidemia, colon cancer, type II diabetes, and atherosclerotic coronary artery disease.
56	ENDCNOT04	Library was constructed using RNA isolated from coronary artery endothelial cell tissue removed from a 3-year-old Caucasian male.



Table 4 (cont.)

Polynucleotide SEQ ID NO:	Library	Library Comment
57	OVARDIT01	Library was constructed using RNA isolated from diseased ovary tissue removed from a 39-year-old Caucasian female during total abdominal hysterectomy, bilateral salpingo-oophorectomy, dilation and curettage, partial colectomy, incidental appendectomy, and temporary colostomy. Pathology indicated the right and left adnexa were extensively involved by endometriosis. Endometriosis also involved the anterior and posterior serosal surfaces of the uterus and the cul-de-sac, mesentery, and muscularis propria of the sigmoid colon. Pathology for the associated tumor tissue indicated multiple (3) intramural, 1 subserosal leiomyomata. Family history included hyperlipidemia, benign hypertension, atherosclerotic coronary artery disease, depressive disorder, brain cancer, and type II diabetes.
58	BRSTTMT02	Library was constructed using RNA isolated from diseased right breast tissue removed from a 46-year-old Caucasian female during a unilateral extended simple mastectomy and open breast biopsy. Pathology indicated mildly proliferative fibrocystic change, including intraductal duct ectasia, papilloma formation, and ductal hyperplasia. Pathology for the associated tumor tissue indicated multifocal ductal carcinoma in situ, both comedo and non-comedo types, nuclear grade 2 with extensive intraductal calcifications. Patient history included deficiency anemia, normal delivery, chronic sinusitis, extrinsic asthma, and kidney infection. Family history included type II diabetes, benign hypertension, cerebrovascular disease, skin cancer, and hyperlipidemia.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value=1.06E-6 Assembled ESTs: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less Full Length sequences: fastx score=100 or greater
BLIMPS	A BLOCKS IMPROVED Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families

Table (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Normalized quality score > GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. <i>supra</i> ; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:
  - 5 a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-23 and SEQ ID NO:25-29,
  - b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-23 and SEQ ID NO:25-29,
  - c) a biologically active fragment of an amino acid sequence selected from the group
  - 10 consisting of SEQ ID NO:1-23 and SEQ ID NO:25-29, and
  - d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-23 and SEQ ID NO:25-29.
- 15 2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-23 and SEQ ID NO:25-29.
3. An isolated polynucleotide encoding a polypeptide of claim 1.
4. An isolated polynucleotide of claim 3 selected from the group consisting of SEQ ID
- 20 NO:30-52 and SEQ ID NO:54-58.
5. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
- 25 6. A cell transformed with a recombinant polynucleotide of claim 5.
7. A transgenic organism comprising a recombinant polynucleotide of claim 5.
8. A method for producing a polypeptide of claim 1, the method comprising:
  - 30 a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and
  - b) recovering the polypeptide so expressed.
- 35

9. An isolated antibody which specifically binds to a polypeptide of claim 1.
10. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:
- 5       a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:30-52 and SEQ ID NO:54-58,
- b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:30-52 and SEQ ID NO:54-58,
- 10       c) a polynucleotide sequence complementary to a),
- d) a polynucleotide sequence complementary to b), and
- e) an RNA equivalent of a)-d).
11. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a
- 15 polynucleotide of claim 10.
12. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 10, the method comprising:
- a) hybridizing the sample with a probe comprising at least 16 contiguous nucleotides
- 20 comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide, and
- b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.
- 25
13. A method of claim 12, wherein the probe comprises at least 30 contiguous nucleotides.
14. A method of claim 12, wherein the probe comprises at least 60 contiguous nucleotides.
- 30
15. A pharmaceutical composition comprising an effective amount of a polypeptide of claim 1 and a pharmaceutically acceptable excipient.
16. A method for treating a disease or condition associated with decreased expression of functional HTMP, comprising administering to a patient in need of such treatment the pharmaceutical
- 35 composition of claim 15.

17. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting agonist activity in the sample.

5

18. A pharmaceutical composition comprising an agonist compound identified by a method of claim 17 and a pharmaceutically acceptable excipient.

19. A method for treating a disease or condition associated with decreased expression of functional HTMP, comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 18.

20. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting antagonist activity in the sample.

21. A pharmaceutical composition comprising an antagonist compound identified by a method of claim 20 and a pharmaceutically acceptable excipient.

20

22. A method for treating a disease or condition associated with overexpression of functional HTMP, comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 21.

23. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 4, the method comprising:

- a) exposing a sample comprising the target polynucleotide to a compound, and
- b) detecting altered expression of the target polynucleotide.

25

PF-0681 PCT

## Sequence Listing

&lt;110&gt; INCYTE PHARMACEUTICALS, INC.

YUE, Henry  
 LAL, Preeti  
 TANG, Y. Tom  
 HILLMAN, Jennifer L.  
 REDDY, Roopa  
 BANDMAN, Olga  
 BAUGHN, Mariah R.  
 Lu, Dyung Aina L.  
 AZIMZAI, Yalda  
 YANG, Junming

&lt;120&gt; HUMAN TRANSMEMBRANE PROTEINS

&lt;130&gt; PF-0681 PCT

&lt;140&gt; To Be Assigned

&lt;141&gt; Herewith

&lt;150&gt; 60/125,537; 60 139,565

&lt;151&gt; 1999-03-22; 1999-06-16

&lt;160&gt; 58

&lt;170&gt; PERL Program

&lt;210&gt; 1

&lt;211&gt; 417

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1316927CD1

&lt;400&gt; 1

Met	Trp	Lys	Gly	Leu	Val	Lys	Arg	Asn	Ala	Ser	Val	Glu	Thr	Val
1				5					10					15
Asp	Asn	Lys	Thr	Ser	Glu	Asp	Val	Thr	Met	Ala	Ala	Ala	Ser	Pro
				20					25					30
Val	Thr	Leu	Thr	Lys	Gly	Thr	Ser	Ala	Ala	His	Leu	Asn	Ser	Met
				35					40					45
Glu	Val	Thr	Thr	Glu	Asp	Thr	Ser	Arg	Thr	Asp	Val	Ser	Glu	Pro
				50					55					60
Ala	Thr	Ser	Gly	Gly	Ala	Ala	Asp	Gly	Val	Thr	Ser	Ile	Ala	Pro
				65					70					75
Thr	Ala	Val	Ala	Ser	Ser	Thr	Thr	Ala	Ala	Ser	Ile	Thr	Thr	Ala
				80					85					90
Ala	Ser	Ser	Met	Thr	Val	Ala	Ser	Ser	Ala	Pro	Thr	Thr	Ala	Ala
				95					100					105
Ser	Ser	Thr	Thr	Val	Ala	Ser	Ile	Ala	Pro	Thr	Thr	Ala	Ala	Ser
				110					115					120
Ser	Met	Thr	Ala	Ala	Ser	Ser	Thr	Pro	Met	Thr	Leu	Ala	Leu	Pro
				125					130					135
Ala	Pro	Thr	Ser	Thr	Ser	Thr	Gly	Arg	Thr	Pro	Ser	Thr	Thr	Ala
				140					145					150
Thr	Gly	His	Pro	Ser	Leu	Ser	Thr	Ala	Leu	Ala	Gln	Val	Pro	Lys
				155					160					165

PF-0681 PCT

```

Ser Ser Ala Leu Pro Arg Thr Ala Thr Leu Ala Thr Leu Ala Thr
170 175 180
Arg Ala Gln Thr Val Ala Thr Thr Ala Asn Thr Ser Ser Pro Met
185 190 195
Ser Thr Arg Pro Ser Pro Ser Lys His Met Pro Ser Asp Thr Ala
200 205 210
Ala Ser Pro Val Pro Pro Met Arg Pro Gln Ala Gln Gly Pro Ile
215 220 225
Ser Gln Val Ser Val Asp Gln Pro Val Val Asn Thr Thr Asn Lys
230 235 240
Ser Thr Pro Met Pro Ser Asn Thr Thr Pro Glu Pro Ala Pro Thr
245 250 255
Pro Thr Val Val Thr Thr Thr Lys Ala Gln Ala Arg Glu Pro Thr
260 265 270
Ala Ser Pro Val Pro Val Pro His Thr Ser Pro Ile Pro Glu Met
275 280 285
Glu Ala Met Ser Pro Thr Thr Gln Pro Ser Pro Met Pro Tyr Thr
290 295 300
Gln Arg Ala Ala Gly Pro Gly Thr Ser Gln Ala Pro Glu Gln Val
305 310 315
Glu Thr Glu Ala Thr Pro Gly Thr Asp Ser Thr Gly Pro Thr Pro
320 325 330
Arg Ser Ser Gly Gly Thr Lys Met Pro Ala Thr Asp Ser Cys Gln
335 340 345
Pro Ser Thr Gln Gly Gln Tyr Met Val Val Thr Thr Glu Pro Leu
350 355 360
Thr Gln Ala Val Val Asp Lys Thr Leu Leu Leu Val Val Leu Leu
365 370 375
Leu Gly Val Thr Leu Phe Ile Thr Val Leu Val Leu Phe Ala Leu
380 385 390
Gln Ala Tyr Glu Ser Tyr Lys Lys Lys Asp Tyr Thr Gln Val Asp
395 400 405
Tyr Leu Ile Asn Gly Met Tyr Ala Asp Ser Glu Met
410 415

```

&lt;210&gt; 2

&lt;211&gt; 316

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1354891CD1

&lt;400&gt; 2

```

Met Pro Lys Asn Phe Trp Ala Gly Val Gly Ile Val Cys Arg Phe
1 5 10 15
Arg Glu Gln Asp Ser Phe Pro Pro Ala Ser Ala Ser Cys Pro Ala
20 25 30
Leu Leu Ala Leu Ser Pro Thr Leu Pro Ser Pro Val Phe Leu Phe
35 40 45
Gln Val Pro Ala Ile Glu Glu Asn Leu Leu Asp Asp Lys His Leu
50 55 60
Leu Lys Pro Trp Asp Ala Lys Lys Leu Ser Ser Ser Ser Ser Arg
65 70 75
Pro Arg Ser Cys Glu Val Pro Gly Ile Asn Ile Phe Pro Ser Pro
80 85 90
Asp Gln Pro Ala Asn Val Pro Val Leu Pro Pro Ala Met Asn Thr
95 100 105
Gly Gly Ser Leu Pro Asp Leu Thr Asn Leu His Phe Pro Pro Pro
110 115 120
Leu Pro Thr Pro Leu Asp Pro Glu Glu Thr Ala Tyr Pro Ser Leu

```



PF-0681 PCT

				125					130					135
Ser	Gly	Gly	Asn	Ser	Thr	Ser	Asn	Leu	Thr	His	Thr	Met	Thr	His
				140					145					150
Leu	Gly	Ile	Ser	Arg	Gly	Met	Gly	Leu	Gly	Pro	Gly	Tyr	Asp	Ala
				155					160					165
Pro	Gly	Leu	His	Ser	Pro	Leu	Ser	His	Pro	Ser	Leu	Gln	Ser	Ser
				170					175					180
Leu	Ser	Asn	Pro	Asn	Leu	Gln	Ala	Ser	Leu	Ser	Ser	Pro	Gln	Pro
				185					190					195
Gln	Leu	Gln	Gly	Ser	His	Ser	His	Pro	Ser	Leu	Pro	Ala	Ser	Ser
				200					205					210
Leu	Ala	Arg	His	Val	Leu	Pro	Thr	Thr	Ser	Leu	Gly	His	Pro	Ser
				215					220					225
Leu	Ser	Ala	Pro	Ala	Leu	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Thr
				230					235					240
Ser	Ser	Pro	Val	Leu	Gly	Ala	Pro	Ser	Tyr	Pro	Ala	Ser	Thr	Pro
				245					250					255
Gly	Ala	Ser	Pro	His	His	Arg	Arg	Val	Pro	Leu	Ser	Pro	Leu	Ser
				260					265					270
Leu	Leu	Ala	Gly	Pro	Ala	Asp	Ala	Arg	Arg	Ser	Gln	Gln	Gln	Leu
				275					280					285
Pro	Lys	Gln	Phe	Ser	Pro	Thr	Met	Ser	Pro	Thr	Leu	Ser	Ser	Ile
				290					295					300
Thr	Gln	Gly	Val	Pro	Trp	Ile	Pro	Val	Asn	Cys	Pro	Leu	Thr	Ser
				305					310					315

Gly

&lt;210&gt; 3

&lt;211&gt; 273

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1415487CD1

&lt;400&gt; 3

Met	Ile	Leu	Ser	Ser	Tyr	Phe	Ile	Asn	Phe	Ile	Tyr	Leu	Ala	Lys
1				5					10					15
Ser	Thr	Lys	Lys	Thr	Met	Leu	Thr	Leu	Thr	Leu	Val	Cys	Ala	Ile
				20					25					30
Thr	Phe	Leu	Leu	Val	Cys	Ser	Gly	Thr	Phe	Phe	Pro	Tyr	Ser	Ser
				35					40					45
Asn	Pro	Ala	Asn	Pro	Lys	Pro	Lys	Arg	Val	Phe	Leu	Gln	His	Met
				50					55					60
Thr	Arg	Thr	Phe	His	Asp	Leu	Glu	Gly	Asn	Ala	Val	Lys	Arg	Asp
				65					70					75
Ser	Gly	Ile	Trp	Ile	Asn	Gly	Phe	Asp	Tyr	Thr	Gly	Ile	Ser	His
				80					85					90
Ile	Thr	Pro	His	Ile	Pro	Glu	Ile	Asn	Asp	Ser	Ile	Arg	Ala	His
				95					100					105
Cys	Glu	Glu	Asn	Ala	Pro	Leu	Cys	Gly	Phe	Pro	Trp	Tyr	Leu	Pro
				110					115					120
Val	His	Phe	Leu	Ile	Arg	Lys	Asn	Trp	Tyr	Leu	Pro	Ala	Pro	Glu
				125					130					135
Val	Ser	Pro	Arg	Asn	Pro	Pro	His	Phe	Arg	Leu	Ile	Ser	Lys	Glu
				140					145					150
Gln	Thr	Pro	Trp	Asp	Ser	Ile	Lys	Leu	Thr	Phe	Glu	Ala	Thr	Gly
				155					160					165
Pro	Ser	His	Met	Ser	Phe	Tyr	Val	Arg	Ala	His	Lys	Gly	Ser	Thr
				170					175					180

PF-0681 PCT

Leu	Ser	Gln	Trp	Ser	Leu	Gly	Asn	Gly	Thr	Pro	Val	Thr	Ser	Lys
				185					190					195
Gly	Gly	Asp	Tyr	Phe	Val	Phe	Tyr	Ser	His	Gly	Leu	Gln	Ala	Ser
				200					205					210
Ala	Trp	Gln	Phe	Trp	Ile	Glu	Val	Gln	Val	Ser	Glu	Glu	His	Pro
				215					220					225
Glu	Gly	Met	Val	Thr	Val	Ala	Ile	Ala	Ala	His	Tyr	Leu	Ser	Gly
				230					235					240
Glu	Asp	Lys	Arg	Ser	Pro	Gln	Leu	Asp	Ala	Leu	Lys	Glu	Lys	Phe
				245					250					255
Pro	Asp	Trp	Thr	Phe	Pro	Ser	Ala	Trp	Val	Cys	Thr	Tyr	Asp	Leu
				260					265					270
Phe	Val	Phe												

&lt;210&gt; 4

&lt;211&gt; 222

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1693184CD1

&lt;400&gt; 4

Met	Cys	Leu	Arg	Leu	Gly	Gly	Leu	Ser	Val	Gly	Asp	Phe	Arg	Lys
1				5					10					15
Val	Leu	Met	Lys	Thr	Gly	Leu	Val	Leu	Val	Val	Leu	Gly	His	Val
				20					25					30
Ser	Phe	Ile	Thr	Ala	Ala	Leu	Phe	His	Gly	Thr	Val	Leu	Arg	Tyr
				35					40					45
Val	Gly	Thr	Pro	Gln	Asp	Ala	Val	Ala	Leu	Gln	Tyr	Cys	Val	Val
				50					55					60
Asn	Ile	Leu	Ser	Val	Thr	Ser	Ala	Ile	Val	Val	Ile	Thr	Ser	Gly
				65					70					75
Ile	Ala	Ala	Ile	Val	Leu	Ser	Arg	Tyr	Leu	Pro	Ser	Thr	Pro	Leu
				80					85					90
Arg	Trp	Thr	Val	Phe	Ser	Ser	Ser	Val	Ala	Cys	Ala	Leu	Leu	Ser
				95					100					105
Leu	Thr	Cys	Ala	Leu	Gly	Leu	Leu	Ala	Ser	Ile	Ala	Met	Thr	Phe
				110					115					120
Ala	Thr	Gln	Gly	Lys	Ala	Leu	Leu	Ala	Ala	Cys	Thr	Phe	Gly	Ser
				125					130					135
Ser	Glu	Leu	Leu	Ala	Leu	Ala	Pro	Asp	Cys	Pro	Phe	Asp	Pro	Thr
				140					145					150
Arg	Ile	Tyr	Ser	Ser	Ser	Leu	Cys	Leu	Trp	Gly	Ile	Ala	Leu	Val
				155					160					165
Leu	Cys	Val	Ala	Glu	Asn	Val	Phe	Ala	Val	Arg	Cys	Ala	Gln	Leu
				170					175					180
Thr	His	Gln	Leu	Leu	Glu	Leu	Arg	Pro	Trp	Trp	Gly	Lys	Ser	Ser
				185					190					195
His	His	Met	Met	Arg	Glu	Asn	Pro	Glu	Leu	Val	Glu	Gly	Arg	Asp
				200					205					210
Leu	Leu	Ser	Cys	Thr	Ser	Ser	Glu	Pro	Leu	Thr	Leu			
				215					220					

&lt;210&gt; 5

&lt;211&gt; 329

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

PF-0681 PCT

&lt;223&gt; Incyte ID No: 1695049CD1

&lt;400&gt; 5

Met	Pro	Gly	Lys	His	Gln	His	Phe	Gln	Glu	Pro	Glu	Val	Gly	Cys	1	5	10	15
Cys	Gly	Lys	Tyr	Phe	Leu	Phe	Gly	Phe	Asn	Ile	Val	Phe	Trp	Val	20	25	30	35
Leu	Gly	Ala	Leu	Phe	Leu	Ala	Ile	Gly	Leu	Trp	Ala	Trp	Gly	Glu	40	45	50	55
Lys	Gly	Val	Leu	Ser	Asn	Ile	Ser	Ala	Leu	Thr	Asp	Leu	Gly	Gly	60	65	70	75
Leu	Asp	Pro	Val	Trp	Leu	Phe	Val	Val	Val	Gly	Gly	Val	Met	Ser	80	85	90	95
Val	Leu	Gly	Phe	Ala	Gly	Cys	Ile	Gly	Ala	Leu	Arg	Glu	Asn	Thr	100	105	110	115
Phe	Leu	Leu	Lys	Phe	Phe	Ser	Val	Phe	Leu	Gly	Leu	Ile	Phe	Phe	120	125	130	135
Leu	Glu	Leu	Ala	Thr	Gly	Ile	Leu	Ala	Phe	Val	Phe	Lys	Asp	Trp	140	145	150	155
Ile	Arg	Asp	Gln	Leu	Asn	Leu	Phe	Ile	Asn	Asn	Asn	Val	Lys	Ala	160	165	170	175
Tyr	Arg	Asp	Asp	Ile	Asp	Leu	Gln	Asn	Leu	Ile	Asp	Phe	Ala	Gln	180	185	190	195
Glu	Tyr	Trp	Ser	Cys	Cys	Gly	Ala	Arg	Gly	Pro	Asn	Asp	Trp	Asn	200	205	210	215
Leu	Asn	Ile	Tyr	Phe	Asn	Cys	Thr	Asp	Leu	Asn	Pro	Ser	Arg	Glu	220	225	230	235
Arg	Cys	Gly	Val	Pro	Phe	Ser	Cys	Cys	Val	Arg	Asp	Pro	Ala	Glu	240	245	250	255
Asp	Val	Leu	Asn	Thr	Gln	Cys	Gly	Tyr	Asp	Val	Arg	Leu	Lys	Leu	260	265	270	275
Glu	Leu	Glu	Gln	Gln	Gly	Phe	Ile	His	Thr	Lys	Gly	Cys	Val	Gly	280	285	290	295
Gln	Phe	Glu	Lys	Trp	Leu	Gln	Asp	Asn	Leu	Ile	Val	Val	Ala	Gly	300	305	310	315
Val	Phe	Met	Gly	Ile	Ala	Leu	Leu	Gln	Ile	Phe	Gly	Ile	Cys	Leu	320	325		
Ala	Gln	Asn	Leu	Val	Ser	Asp	Ile	Lys	Ala	Val	Lys	Ala	Asn	Trp				
Ser	Lys	Trp	Asn	Asp	Asp	Phe	Glu	Asn	His	Trp	Leu	Thr	Pro	Thr				
Ile	Ser	Glu	Val	Leu	Ser	Thr	Ala	Gly	Pro	Gln	Gln	Asn	Ser	Leu				
Thr	Gly	Ala	Pro	Gly	Pro	Ala	Pro	Pro	Ser	Arg	His	Val	Phe	Phe				
Gly	Leu	Gly	Gly	Leu	Tyr	Pro	Glu	Pro	Thr	Phe	Lys	Asn	Trp					

&lt;210&gt; 6

&lt;211&gt; 351

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1802448CD1

&lt;400&gt; 6

Met	Ala	Leu	Thr	Gly	Ala	Ser	Asp	Pro	Ser	Ala	Glu	Ala	Glu	Ala	1	5	10	15
Asn	Gly	Glu	Lys	Pro	Phe	Leu	Leu	Arg	Ala	Leu	Gln	Ile	Ala	Leu	20	25	30	

PF-0681 PCT

```

Val Val Ser Leu Tyr Trp Val Thr Ser Ile Ser Met Val Phe Leu
    35          40          45
Asn Lys Tyr Leu Leu Asp Ser Pro Ser Leu Arg Leu Asp Thr Pro
    50          55          60
Ile Phe Val Thr Phe Tyr Gln Cys Leu Val Thr Thr Leu Leu Cys
    65          70          75
Lys Gly Leu Ser Ala Leu Ala Ala Cys Cys Pro Gly Ala Val Asp
    80          85          90
Phe Pro Ser Leu Arg Leu Asp Leu Arg Val Ala Arg Ser Val Leu
    95         100         105
Pro Leu Ser Val Val Phe Ile Gly Met Ile Thr Phe Asn Asn Leu
   110         115         120
Cys Leu Lys Tyr Val Gly Val Ala Phe Tyr Asn Val Gly Arg Ser
   125         130         135
Leu Thr Thr Val Phe Asn Val Leu Leu Ser Tyr Leu Leu Leu Lys
   140         145         150
Gln Thr Thr Ser Phe Tyr Ala Leu Leu Thr Cys Gly Ile Ile Ile
   155         160         165
Gly Gly Phe Trp Leu Gly Val Asp Gln Glu Gly Ala Glu Gly Thr
   170         175         180
Leu Ser Trp Leu Gly Thr Val Phe Gly Val Leu Ala Ser Leu Cys
   185         190         195
Val Ser Leu Asn Ala Ile Tyr Thr Thr Lys Val Leu Pro Ala Val
   200         205         210
Asp Gly Ser Ile Trp Arg Leu Thr Phe Tyr Asn Asn Val Asn Ala
   215         220         225
Cys Ile Leu Phe Leu Pro Leu Leu Leu Leu Leu Gly Glu Leu Gln
   230         235         240
Ala Leu Arg Asp Phe Ala Gln Leu Gly Ser Ala His Phe Trp Gly
   245         250         255
Met Met Thr Leu Gly Gly Leu Phe Gly Phe Ala Ile Gly Tyr Val
   260         265         270
Thr Gly Leu Gln Ile Lys Phe Thr Ser Pro Leu Thr His Asn Val
   275         280         285
Ser Gly Thr Ala Lys Ala Cys Ala Gln Thr Val Leu Ala Val Leu
   290         295         300
Tyr Tyr Glu Glu Thr Lys Ser Phe Leu Trp Trp Thr Ser Asn Met
   305         310         315
Met Val Leu Gly Gly Ser Ser Ala Tyr Thr Trp Val Arg Gly Trp
   320         325         330
Glu Met Lys Lys Thr Pro Glu Glu Pro Ser Pro Lys Asp Ser Glu
   335         340         345
Lys Ser Ala Met Gly Val
    350

```

&lt;210&gt; 7

&lt;211&gt; 489

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2057214CD1

&lt;400&gt; 7

```

Met Glu Ala Pro Asp Tyr Glu Val Leu Ser Val Arg Glu Gln Leu
    1          5          10          15
Phe His Glu Arg Ile Arg Glu Cys Ile Ile Ser Thr Leu Leu Phe
    20          25          30
Ala Thr Leu Tyr Ile Leu Cys His Ile Phe Leu Thr Arg Phe Lys
    35          40          45
Lys Pro Ala Glu Phe Thr Thr Val Asp Asp Glu Asp Ala Thr Val

```

PF-0681 PCT

	50		55		60
Asn Lys Ile Ala	Leu Glu Leu Cys Thr	Phe Thr Leu Ala Ile			
	65		70		75
Leu Gly Ala Val	Leu Leu Leu Pro Phe	Ser Ile Ile Ser Asn Glu			
	80		85		90
Val Leu Leu Ser	Leu Pro Arg Asn Tyr	Tyr Ile Gln Trp Leu Asn			
	95		100		105
Gly Ser Leu Ile	His Gly Leu Trp Asn	Leu Val Phe Leu Phe Ser			
	110		115		120
Asn Leu Ser Leu	Ile Phe Leu Met Pro	Phe Ala Tyr Phe Phe Thr			
	125		130		135
Glu Ser Glu Gly	Phe Ala Gly Ser Arg	Lys Gly Val Leu Gly Arg			
	140		145		150
Val Tyr Glu Thr	Val Val Met Leu Met	Leu Leu Thr Leu Leu Val			
	155		160		165
Leu Gly Met Val	Trp Val Ala Ser Ala	Ile Val Asp Lys Asn Lys			
	170		175		180
Ala Asn Arg Glu	Ser Leu Tyr Asp Phe	Trp Glu Tyr Tyr Leu Pro			
	185		190		195
Tyr Leu Tyr Ser	Cys Ile Ser Phe Leu	Gly Val Leu Leu Leu Leu			
	200		205		210
Val Cys Thr Pro	Leu Gly Leu Ala Arg	Met Phe Ser Val Thr Gly			
	215		220		225
Lys Leu Leu Val	Lys Pro Arg Leu Leu	Glu Asp Leu Glu Glu Gln			
	230		235		240
Leu Tyr Cys Ser	Ala Phe Glu Glu Ala	Ala Leu Thr Arg Arg Ile			
	245		250		255
Cys Asn Pro Thr	Ser Cys Trp Leu Pro	Leu Asp Met Glu Leu Leu			
	260		265		270
His Arg Gln Val	Leu Ala Leu Gln Thr	Gln Arg Val Leu Leu Glu			
	275		280		285
Lys Arg Arg Lys	Ala Ser Ala Trp Gln	Arg Asn Leu Gly Tyr Pro			
	290		295		300
Leu Ala Met Leu	Cys Leu Leu Val Leu	Thr Gly Leu Ser Val Leu			
	305		310		315
Ile Val Ala Ile	His Ile Leu Glu Leu	Leu Ile Asp Glu Ala Ala			
	320		325		330
Met Pro Arg Gly	Met Gln Gly Thr Ser	Leu Gly Gln Val Ser Phe			
	335		340		345
Ser Lys Leu Gly	Ser Phe Gly Ala Val	Ile Gln Val Val Leu Ile			
	350		355		360
Phe Tyr Leu Met	Val Ser Ser Val Val	Gly Phe Tyr Ser Ser Pro			
	365		370		375
Leu Phe Arg Ser	Leu Arg Pro Arg Trp	His Asp Thr Ala Met Thr			
	380		385		390
Gln Ile Ile Gly	Asn Cys Val Cys Leu	Leu Val Leu Ser Ser Ala			
	395		400		405
Leu Pro Val Phe	Ser Arg Thr Leu Gly	Leu Thr Arg Phe Asp Leu			
	410		415		420
Leu Gly Asp Phe	Gly Arg Phe Asn Trp	Leu Gly Asn Phe Tyr Ile			
	425		430		435
Val Phe Leu Tyr	Asn Ala Ala Phe Ala	Gly Leu Thr Thr Leu Cys			
	440		445		450
Leu Val Lys Thr	Phe Thr Ala Ala Val	Arg Ala Glu Leu Ile Arg			
	455		460		465
Ala Phe Gly Leu	Asp Arg Leu Pro Leu	Pro Val Ser Gly Phe Pro			
	470		475		480
Gln Ala Ser Arg	Lys Thr Gln His Gln				
	485				

&lt;210&gt; 8

&lt;211&gt; 291

PF-0681 PCT

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2448177CD1

&lt;400&gt; 8

Met	Val	Trp	Lys	Lys	Leu	Gly	Ser	Arg	Asn	Phe	Ser	Ser	Cys	Pro
1				5					10					15
Ser	Gly	Ser	Ile	Gln	Trp	Ile	Trp	Asp	Val	Leu	Gly	Glu	Cys	Ala
				20					25					30
Gln	Asp	Gly	Trp	Asp	Glu	Ala	Ser	Val	Gly	Leu	Gly	Leu	Ile	Ser
				35					40					45
Ile	Leu	Cys	Phe	Ala	Ala	Ser	Thr	Phe	Pro	Gln	Phe	Ile	Lys	Ala
				50					55					60
Tyr	Lys	Thr	Gly	Asn	Met	Asp	Gln	Ala	Leu	Ser	Leu	Trp	Phe	Leu
				65					70					75
Leu	Gly	Trp	Ile	Gly	Gly	Asp	Ser	Cys	Asn	Leu	Ile	Gly	Ser	Phe
				80					85					90
Leu	Ala	Asp	Gln	Leu	Pro	Leu	Gln	Thr	Tyr	Thr	Ala	Val	Tyr	Tyr
				95					100					105
Val	Leu	Ala	Asp	Leu	Val	Met	Leu	Thr	Leu	Tyr	Phe	Tyr	Tyr	Lys
				110					115					120
Phe	Arg	Thr	Arg	Pro	Ser	Leu	Leu	Ser	Ala	Pro	Ile	Asn	Ser	Val
				125					130					135
Leu	Leu	Phe	Leu	Met	Gly	Met	Ala	Cys	Ala	Thr	Pro	Leu	Leu	Ser
				140					145					150
Ala	Ala	Gly	Pro	Val	Ala	Ala	Pro	Arg	Glu	Ala	Phe	Arg	Gly	Arg
				155					160					165
Ala	Leu	Leu	Ser	Val	Glu	Ser	Gly	Ser	Lys	Pro	Phe	Thr	Arg	Gln
				170					175					180
Glu	Val	Ile	Gly	Phe	Val	Ile	Gly	Ser	Ile	Ser	Ser	Val	Leu	Tyr
				185					190					195
Leu	Leu	Ser	Arg	Leu	Pro	Gln	Ile	Arg	Thr	Asn	Phe	Leu	Arg	Lys
				200					205					210
Ser	Thr	Gln	Gly	Ile	Ser	Tyr	Ser	Leu	Phe	Ala	Leu	Val	Met	Leu
				215					220					225
Gly	Asn	Thr	Leu	Tyr	Gly	Leu	Ser	Val	Leu	Leu	Lys	Asn	Pro	Glu
				230					235					240
Glu	Gly	Gln	Ser	Glu	Gly	Ser	Tyr	Leu	Leu	His	His	Leu	Pro	Trp
				245					250					255
Leu	Val	Gly	Ser	Leu	Gly	Val	Leu	Leu	Leu	Asp	Thr	Ile	Ile	Ser
				260					265					270
Ile	Gln	Phe	Leu	Val	Tyr	Arg	Arg	Ser	Thr	Ala	Ala	Ser	Glu	Leu
				275					280					285
Glu	Pro	Leu	Leu	Pro	Ser									
				290										

&lt;210&gt; 9

&lt;211&gt; 172

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2741701CD1

&lt;400&gt; 9

Met	Ser	Ser	Ser	Gly	Gly	Ala	Pro	Gly	Ala	Ser	Ala	Ser	Ser	Ala
1				5					10					15
Pro	Pro	Ala	Gln	Glu	Glu	Gly	Met	Thr	Trp	Trp	Tyr	Arg	Trp	Leu

PF-0681 PCT

	20		25		30
Cys Arg Leu Ser Gly Val Leu Gly Ala Val Ser Cys Ala Ile Ser					
	35		40		45
Gly Leu Phe Asn Cys Ile Thr Ile His Pro Leu Asn Ile Ala Ala					
	50		55		60
Gly Val Trp Met Ile Met Asn Ala Phe Ile Leu Leu Leu Cys Glu					
	65		70		75
Ala Pro Phe Cys Cys Gln Phe Ile Glu Phe Ala Asn Thr Val Ala					
	80		85		90
Glu Lys Val Asp Arg Leu Arg Ser Trp Gln Lys Ala Val Phe Tyr					
	95		100		105
Cys Gly Met Ala Val Val Pro Ile Val Ile Ser Leu Thr Leu Thr					
	110		115		120
Thr Leu Leu Gly Asn Ala Ile Ala Phe Ala Thr Gly Val Leu Tyr					
	125		130		135
Gly Leu Ser Ala Leu Gly Lys Lys Gly Asp Ala Ile Ser Tyr Ala					
	140		145		150
Arg Ile Gln Gln Gln Arg Gln Gln Ala Asp Glu Glu Lys Leu Ala					
	155		160		165
Glu Thr Leu Glu Gly Glu Leu					
	170				

&lt;210&gt; 10

&lt;211&gt; 155

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3487228CD1

&lt;400&gt; 10

Met Ala Gly Glu Ile Thr Glu Thr Gly Glu Leu Tyr Ser Ser Tyr					
1	5		10		15
Val Gly Leu Val Tyr Met Phe Asn Leu Ile Val Gly Thr Gly Ala					
	20		25		30
Leu Thr Met Pro Lys Ala Phe Ala Thr Ala Gly Trp Leu Val Ser					
	35		40		45
Leu Val Leu Leu Val Phe Leu Gly Phe Met Ser Phe Met Thr Thr					
	50		55		60
Thr Phe Val Ile Glu Ala Met Ala Ala Ala Asn Ala Gln Leu His					
	65		70		75
Trp Lys Arg Met Glu Asn Leu Lys Glu Glu Glu Asp Asp Asp Ser					
	80		85		90
Ser Thr Ala Ser Asp Ser Asp Val Leu Ile Arg Asp Asn Tyr Glu					
	95		100		105
Arg Ala Glu Lys Arg Pro Ile Leu Ser Val Arg Lys Ser Trp Gly					
	110		115		120
Ser Arg Gly Pro Ala Leu Pro Ser Gly Pro Gln Gly Ser Leu Ser					
	125		130		135
Ser Trp Arg Gly Pro Trp Arg Gly Ser Pro Gln Ile Leu Ala Phe					
	140		145		150
Pro Ala Pro Gly Phe					
	155				

&lt;210&gt; 11

&lt;211&gt; 578

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3671426CD1

PF-0681 PCT

&lt;400&gt; 11

Met	Leu	Cys	Ser	Leu	Leu	Leu	Cys	Glu	Cys	Leu	Leu	Leu	Val	Ala
1				5					10					15
Gly	Tyr	Ala	His	Asp	Asp	Asp	Trp	Ile	Asp	Pro	Thr	Asp	Met	Leu
				20					25					30
Asn	Tyr	Asp	Ala	Ala	Ser	Gly	Thr	Met	Arg	Lys	Ser	Gln	Ala	Lys
				35					40					45
Tyr	Gly	Ile	Ser	Gly	Glu	Lys	Asp	Val	Ser	Pro	Asp	Leu	Ser	Cys
				50					55					60
Ala	Asp	Glu	Ile	Ser	Glu	Cys	Tyr	His	Lys	Leu	Asp	Ser	Leu	Thr
				65					70					75
Tyr	Lys	Ile	Asp	Glu	Cys	Glu	Lys	Lys	Lys	Arg	Glu	Asp	Tyr	Glu
				80					85					90
Ser	Gln	Ser	Asn	Pro	Val	Phe	Arg	Arg	Tyr	Leu	Asn	Lys	Ile	Leu
				95					100					105
Ile	Glu	Ala	Gly	Lys	Leu	Gly	Leu	Pro	Asp	Glu	Asn	Lys	Gly	Asp
				110					115					120
Met	His	Tyr	Asp	Ala	Glu	Ile	Ile	Leu	Lys	Arg	Glu	Thr	Leu	Leu
				125					130					135
Glu	Ile	Gln	Lys	Phe	Leu	Asn	Gly	Glu	Asp	Trp	Lys	Pro	Gly	Ala
				140					145					150
Leu	Asp	Asp	Ala	Leu	Ser	Asp	Ile	Leu	Ile	Asn	Phe	Lys	Phe	His
				155					160					165
Asp	Phe	Glu	Thr	Trp	Lys	Trp	Arg	Phe	Glu	Asp	Ser	Phe	Gly	Val
				170					175					180
Asp	Pro	Tyr	Asn	Val	Leu	Met	Val	Leu	Leu	Cys	Leu	Leu	Cys	Ile
				185					190					195
Val	Val	Leu	Val	Ala	Thr	Glu	Leu	Trp	Thr	Tyr	Val	Arg	Trp	Tyr
				200					205					210
Thr	Gln	Leu	Arg	Arg	Val	Leu	Ile	Ile	Ser	Phe	Leu	Phe	Ser	Leu
				215					220					225
Gly	Trp	Asn	Trp	Met	Tyr	Leu	Tyr	Lys	Leu	Ala	Phe	Ala	Gln	His
				230					235					240
Gln	Ala	Glu	Val	Ala	Lys	Met	Glu	Pro	Leu	Asn	Asn	Val	Cys	Ala
				245					250					255
Lys	Lys	Met	Asp	Trp	Thr	Gly	Ser	Ile	Trp	Glu	Trp	Phe	Arg	Ser
				260					265					270
Ser	Trp	Thr	Tyr	Lys	Asp	Asp	Pro	Cys	Gln	Lys	Tyr	Tyr	Glu	Leu
				275					280					285
Leu	Leu	Val	Asn	Pro	Ile	Trp	Leu	Val	Pro	Pro	Thr	Lys	Ala	Leu
				290					295					300
Ala	Val	Thr	Phe	Thr	Thr	Phe	Val	Thr	Glu	Pro	Leu	Lys	His	Ile
				305					310					315
Gly	Lys	Gly	Thr	Gly	Glu	Phe	Ile	Lys	Ala	Leu	Met	Lys	Glu	Ile
				320					325					330
Pro	Ala	Leu	Leu	His	Leu	Pro	Val	Leu	Ile	Ile	Met	Ala	Leu	Ala
				335					340					345
Ile	Leu	Ser	Phe	Cys	Tyr	Gly	Ala	Gly	Lys	Ser	Val	His	Val	Leu
				350					355					360
Arg	His	Ile	Gly	Gly	Pro	Glu	Arg	Glu	Pro	Pro	Gln	Ala	Leu	Trp
				365					370					375
Pro	Arg	Asp	Arg	Arg	Arg	Gln	Glu	Glu	Ile	Asp	Tyr	Arg	Pro	Asp
				380					385					390
Gly	Gly	Ala	Gly	Asp	Ala	Asp	Phe	His	Tyr	Arg	Gly	Gln	Met	Gly
				395					400					405
Pro	Thr	Glu	Gln	Gly	Pro	Tyr	Ala	Lys	Thr	Tyr	Glu	Gly	Arg	Arg
				410					415					420
Glu	Ile	Leu	Arg	Glu	Arg	Asp	Val	Asp	Leu	Arg	Phe	Gln	Thr	Gly
				425					430					435
Asn	Lys	Ser	Pro	Glu	Val	Leu	Arg	Ala	Phe	Asp	Val	Pro	Asp	Ala
				440					445					450



PF-0681 PCT

Glu	Ala	Arg	Glu	His	Pro	Thr	Val	Val	Pro	Ser	His	Lys	Ser	Pro
				455					460					465
Val	Leu	Asp	Thr	Lys	Pro	Lys	Glu	Thr	Gly	Gly	Ile	Leu	Gly	Lys
				470					475					480
Ala	His	Arg	Lys	Lys	Ala	Val	Leu	Lys	Ala	Ala	Ser	Arg	Pro	Ser
				485					490					495
Leu	Ser	Leu	Ala	Lys	Thr	His	Gln	Gly	Ile	Gln	Lys	Val	His	Pro
				500					505					510
Gln	Arg	Lys	Arg	Pro	Ser	Ser	Ser	Leu	Lys	Pro	Gln	Ala	Ala	Gln
				515					520					525
Thr	Lys	Ala	Ala	His	Thr	Ala	Pro	Gln	Glu	Val	Trp	Leu	Asp	His
				530					535					540
Val	Asp	Arg	Ile	Arg	Ser	Ala	Ala	Pro	Val	Ala	Arg	Gly	Thr	Pro
				545					550					555
Ala	Gln	Thr	Thr	Ala	Ser	Ser	Leu	Leu	Arg	Ala	Leu	Tyr	Pro	Phe
				560					565					570
Gly	Asp	Glu	Val	Tyr	Phe	Asp	Ser							
				575										

&lt;210&gt; 12

&lt;211&gt; 313

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 4020882CD1

&lt;400&gt; 12

Met	Asp	Gly	Ile	Ile	Glu	Gln	Lys	Ser	Met	Leu	Val	His	Ser	Lys
1				5					10					15
Ile	Ser	Asp	Ala	Gly	Lys	Arg	Asn	Gly	Leu	Ile	Asn	Thr	Arg	Asn
				20					25					30
Leu	Met	Ala	Glu	Ser	Arg	Asp	Gly	Leu	Val	Ser	Val	Tyr	Pro	Ala
				35					40					45
Pro	Gln	Tyr	Gln	Ser	His	Arg	Val	Gly	Ala	Ser	Thr	Val	Pro	Ala
				50					55					60
Ser	Leu	Asp	Ser	Ser	Arg	Ser	Glu	Pro	Met	Gln	Gln	Leu	Leu	Asp
				65					70					75
Pro	Asn	Thr	Leu	Gln	Gln	Ser	Val	Glu	Ser	Arg	Tyr	Arg	Pro	Asn
				80					85					90
Ile	Ile	Leu	Tyr	Ser	Glu	Gly	Val	Leu	Arg	Ser	Trp	Gly	Asp	Gly
				95					100					105
Val	Ala	Ala	Asp	Cys	Cys	Glu	Thr	Thr	Phe	Ile	Glu	Asp	Arg	Ser
				110					115					120
Pro	Thr	Lys	Asp	Ser	Leu	Glu	Tyr	Pro	Asp	Gly	Lys	Phe	Ile	Asp
				125					130					135
Leu	Ser	Ala	Asp	Asp	Ile	Lys	Ile	His	Thr	Leu	Ser	Tyr	Asp	Val
				140					145					150
Glu	Glu	Glu	Glu	Glu	Phe	Gln	Glu	Leu	Glu	Ser	Asp	Tyr	Ser	Ser
				155					160					165
Asp	Thr	Glu	Ser	Glu	Asp	Asn	Phe	Leu	Met	Met	Pro	Pro	Arg	Asp
				170					175					180
His	Leu	Gly	Leu	Ser	Val	Phe	Ser	Met	Leu	Cys	Cys	Phe	Trp	Pro
				185					190					195
Leu	Gly	Ile	Ala	Ala	Phe	Tyr	Leu	Ser	His	Glu	Thr	Asn	Lys	Ala
				200					205					210
Val	Ala	Lys	Gly	Asp	Leu	His	Gln	Ala	Ser	Thr	Ser	Ser	Arg	Arg
				215					220					225
Ala	Leu	Phe	Leu	Ala	Val	Leu	Ser	Ile	Thr	Ile	Gly	Thr	Gly	Val
				230					235					240
Tyr	Val	Gly	Val	Ala	Trp	Pro	Trp	Pro	Ser	Ser	Pro	Thr	Ser	Pro

Arg	Thr	Thr	Thr	Cys	Glu	Leu	Pro	Ala	Asn	Gly	Gly	Gly	Ala	Pro
Gly	Ala	Arg	Ser	Val	Trp	Thr	Trp	Arg	Lys	Gln	Ala	Tyr	Arg	Met
Met	Leu	Tyr	Ser	Thr	Asn	Asp	Cys	Gln	Met	Met	Pro	Arg	Ser	Pro
Gly	Ile	Ser	Tyr	Pro	Trp	Ile	Tyr	Phe	Val	Phe	Ile	Leu		

```
<220>  
<221> misc_feature  
<223> Incyte ID No: 4069777CD1
```

<400> 13															
Met	Met	Ser	Arg	Asp	Val	Ile	Ala	Leu	Val	Ser	Gln	Val	Leu	Pro	15
1				5					10						
Val	Tyr	Ser	Val	Phe	His	Val	Phe	Glu	Ala	Ile	Cys	Cys	Val	Tyr	30
				20					25						
Gly	Gly	Val	Leu	Arg	Gly	Thr	Gly	Lys	Gln	Ala	Phe	Gly	Ala	Ala	45
				35					40						
Val	Asn	Ala	Ile	Thr	Tyr	Tyr	Ile	Ile	Gly	Leu	Pro	Leu	Gly	Ile	60
				50					55						
Leu	Leu	Thr	Phe	Val	Val	Arg	Met	Arg	Ile	Met	Gly	Leu	Trp	Leu	75
				65					70						
Gly	Met	Leu	Ala	Cys	Val	Phe	Leu	Ala	Thr	Ala	Ala	Phe	Val	Ala	90
				80					85						
Tyr	Thr	Ala	Arg	Leu	Asp	Trp	Lys	Leu	Ala	Ala	Glu	Glu	Ala	Lys	105
				95					100						
Lys	His	Ser	Gly	Arg	Gln	Gln	Gln	Gln	Arg	Ala	Glu	Ser	Thr	Ala	120
				110					115						
Thr	Arg	Pro	Gly	Pro	Glu	Lys	Ala	Val	Leu	Ser	Ser	Val	Ala	Thr	135
				125					130						
Gly	Ser	Ser	Pro	Gly	Ile	Thr	Leu	Thr	Thr	Tyr	Ser	Arg	Ser	Glu	150
				140					145						
Cys	His	Val	Asp	Phe	Phe	Arg	Thr	Pro	Glu	Glu	Ala	His	Ala	Leu	165
				155					160						
Ser	Ala	Pro	Thr	Ser	Arg	Leu	Ser	Val	Lys	Gln	Leu	Val	Ile	Arg	180
				170					175						
Arg	Gly	Ala	Ala	Leu	Gly	Ala	Ala	Ser	Ala	Thr	Leu	Met	Val	Gly	195
				185					190						
Leu	Thr	Val	Arg	Ile	Leu	Ala	Thr	Arg	His						
				200					205						

```
<220>
<221> misc_feature
<223> Incyte ID No: 5309830CD1
```

```

<400> 14
Met Ala Trp Thr Lys Tyr Gln Leu Phe Leu Ala Gly Leu Met Leu
  1          5          10          15
Val Thr Gly Ser Ile Asn Thr Leu Ser Ala Lys Trp Ala Asp Asn
      20          25          30

```

PF-0681 PCT

Phe Met Ala Glu Gly Cys Gly Gly Ser Lys Glu His Ser Phe Gln  
 35 40 45  
 His Pro Phe Leu Gln Ala Val Gly Met Phe Leu Gly Glu Phe Ser  
 50 55 60  
 Cys Leu Ala Ala Phe Tyr Leu Leu Arg Cys Arg Ala Ala Gly Gln  
 65 70 75  
 Ser Asp Ser Ser Val Asp Pro Gln Gln Pro Phe Asn Pro Leu Leu  
 80 85 90  
 Phe Leu Pro Pro Ala Leu Cys Asp Met Thr Gly Thr Ser Leu Met  
 95 100 105  
 Tyr Val Ala Leu Asn Met Thr Ser Ala Ser Ser Phe Gln Met Leu  
 110 115 120  
 Arg Gly Ala Val Ile Ile Phe Thr Gly Leu Phe Ser Val Ala Phe  
 125 130 135  
 Leu Gly Arg Arg Leu Val Leu Ser Gln Trp Leu Gly Ile Leu Ala  
 140 145 150  
 Thr Ile Ala Gly Leu Val Val Val Gly Leu Ala Asp Leu Leu Ser  
 155 160 165  
 Lys His Asp Ser Gln His Lys Leu Ser Glu Val Ile Thr Gly Asp  
 170 175 180  
 Leu Leu Ile Ile Met Ala Gln Ile Ile Val Ala Ile Gln Met Val  
 185 190 195  
 Leu Glu Glu Lys Phe Val Tyr Lys His Asn Val His Pro Leu Arg  
 200 205 210  
 Ala Val Gly Thr Glu Gly Leu Phe Gly Phe Val Ile Leu Ser Leu  
 215 220 225  
 Leu Leu Val Pro Met Tyr Tyr Ile Pro Ala Gly Ser Phe Ser Gly  
 230 235 240  
 Asn Pro Arg Gly Thr Leu Glu Asp Ala Leu Asp Ala Phe Cys Gln  
 245 250 255  
 Val Gly Gln Gln Pro Leu Ile Ala Val Ala Leu Leu Gly Asn Ile  
 260 265 270  
 Ser Ser Ile Ala Phe Phe Asn Phe Ala Gly Ile Ser Val Thr Lys  
 275 280 285  
 Glu Leu Ser Ala Thr Thr Arg Met Val Leu Asp Ser Leu Arg Thr  
 290 295 300  
 Val Val Ile Trp Ala Leu Ser Leu Ala Leu Gly Trp Glu Ala Phe  
 305 310 315  
 His Ala Leu Gln Ile Leu Gly Phe Leu Ile Leu Leu Ile Gly Thr  
 320 325 330  
 Ala Leu Tyr Asn Gly Leu His Arg Pro Leu Leu Gly Arg Leu Ser  
 335 340 345  
 Arg Gly Arg Pro Leu Ala Glu Glu Ser Glu Gln Glu Arg Leu Leu  
 350 355 360  
 Gly Gly Thr Arg Thr Pro Ile Asn Asp Ala Ser  
 365 370

&lt;210&gt; 15

&lt;211&gt; 374

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 306263CD1

&lt;400&gt; 15

Met Ala Tyr Ile Glu Gln Arg Arg Ile Ser His Glu Gly Ser Pro  
 1 5 10 15  
 Val Lys Pro Val Ala Ile Arg Glu Phe Gln Lys Thr Glu Asp Met  
 20 25 30  
 Arg Arg Tyr Leu His Gln Asn Arg Val Pro Ala Glu Pro Ser Ser

PF-0681 PCT

				35					40					45
Leu	Leu	Ser	Leu	Ser	Ala	Ser	His	Asn	Gln	Leu	Ser	His	Thr	Asp
				50					55					60
Leu	Glu	Leu	His	Gln	Arg	Arg	Glu	Gln	Leu	Val	Glu	Arg	Thr	Arg
				65					70					75
Arg	Glu	Ala	Gln	Leu	Ala	Ala	Leu	Gln	Tyr	Glu	Glu	Glu	Lys	Ile
				80					85					90
Arg	Thr	Lys	Gln	Ile	Gln	Arg	Asp	Ala	Val	Leu	Asp	Phe	Val	Lys
				95					100					105
Gln	Lys	Ala	Ser	Gln	Ser	Pro	Gln	Lys	Gln	His	Pro	Leu	Leu	Asp
				110					115					120
Gly	Val	Asp	Gly	Glu	Cys	Pro	Phe	Pro	Ser	Arg	Arg	Ser	Gln	His
				125					130					135
Thr	Asp	Asp	Ser	Ala	Leu	Cys	Met	Ser	Asp	Asp	Arg	Pro	Asn	Ala
				140					145					150
Leu	Leu	Ser	Ser	Pro	Ala	Thr	Glu	Thr	Val	His	His	Ser	Pro	Ala
				155					160					165
Tyr	Ser	Phe	Pro	Ala	Ala	Ile	Gln	Arg	Asn	Gln	Pro	Gln	Arg	Pro
				170					175					180
Glu	Ser	Phe	Leu	Phe	Arg	Ala	Gly	Val	Arg	Ala	Glu	Thr	Asn	Lys
				185					190					195
Gly	His	Ala	Ser	Pro	Leu	Pro	Pro	Ser	Ala	Ala	Pro	Thr	Thr	Asp
				200					205					210
Ser	Thr	Asp	Ser	Ile	Thr	Gly	Gln	Asn	Ser	Arg	Gln	Arg	Glu	Glu
				215					220					225
Glu	Leu	Glu	Leu	Ile	Asp	Gln	Leu	Arg	Lys	His	Ile	Glu	Tyr	Arg
				230					235					240
Leu	Lys	Val	Ser	Leu	Pro	Cys	Asp	Leu	Gly	Ala	Ala	Leu	Thr	Asp
				245					250					255
Gly	Val	Val	Leu	Cys	His	Leu	Ala	Asn	His	Val	Arg	Pro	Arg	Ser
				260					265					270
Val	Pro	Ser	Ile	His	Val	Pro	Ser	Pro	Ala	Val	Pro	Lys	Leu	Thr
				275					280					285
Met	Ala	Lys	Cys	Arg	Arg	Asn	Val	Glu	Asn	Phe	Leu	Glu	Ala	Cys
				290					295					300
Arg	Lys	Ile	Gly	Val	Pro	Gln	Asp	Asn	Leu	Cys	Ser	Pro	Ser	Asp
				305					310					315
Ile	Leu	Gln	Leu	Asn	Leu	Ser	Val	Lys	Arg	Thr	Val	Glu	Thr	Leu
				320					325					330
Leu	Ser	Leu	Gly	Ala	His	Ser	Glu	Glu	Ser	Ser	Phe	Val	Cys	Leu
				335					340					345
Ser	Leu	Gln	Leu	Leu	Gly	Phe	Val	Ala	Phe	Tyr	Cys	Thr	Val	Met
				350					355					360
Leu	Thr	Leu	Cys	Val	Leu	Tyr	Tyr	Trp	Leu	Phe	Pro	Ala	Arg	
				365					370					

&lt;210&gt; 16

&lt;211&gt; 183

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 483751CD1

&lt;400&gt; 16

Met	Glu	Asn	Gly	Ala	Val	Tyr	Ser	Pro	Thr	Thr	Glu	Glu	Asp	Pro
1				5					10					15
Gly	Pro	Ala	Arg	Gly	Pro	Arg	Ser	Gly	Leu	Ala	Ala	Tyr	Phe	Phe
				20					25					30
Met	Gly	Arg	Leu	Pro	Leu	Leu	Arg	Arg	Val	Leu	Lys	Gly	Leu	Gln
				35					40					45

PF-0681 PCT

```

Leu Leu Leu Ser Leu Leu Ala Phe Ile Cys Glu Glu Val Val Ser
      50      55      60
Gln Cys Thr Leu Cys Gly Gly Leu Tyr Phe Phe Glu Phe Val Ser
      65      70      75
Cys Ser Ala Phe Leu Leu Ser Leu Leu Ile Leu Ile Val Tyr Cys
      80      85      90
Thr Pro Phe Tyr Glu Arg Val Asp Thr Thr Lys Val Lys Ser Ser
      95     100     105
Asp Phe Tyr Ile Thr Leu Gly Thr Gly Cys Val Phe Leu Leu Ala
     110     115     120
Ser Ile Ile Phe Val Ser Thr His Asp Arg Thr Ser Ala Glu Ile
     125     130     135
Ala Ala Ile Val Phe Gly Phe Ile Ala Ser Phe Met Phe Leu Leu
     140     145     150
Asp Phe Ile Thr Met Leu Tyr Glu Lys Arg Gln Glu Ser Gln Leu
     155     160     165
Arg Lys Pro Glu Asn Thr Thr Arg Ala Glu Ala Leu Thr Glu Pro
     170     175     180
Leu Asn Ala

```

&lt;210&gt; 17

&lt;211&gt; 190

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 727332CD1

&lt;400&gt; 17

```

Met Asp Thr Thr Pro Leu Ile Thr Glu His Val Ile Asn Thr Thr
  1      5      10
Gln Asp Pro Cys Ser Trp Val Asp Lys Leu Ser Thr Val His His
  20     25     30
Arg Ile Val Gly Cys Ser Leu Ala Val Ile Ser Gly Val Leu Tyr
  35     40     45
Gly Ser Thr Phe Val Pro Ile Ile Tyr Ile Lys Asp His Ser Lys
  50     55     60
Arg Asn Asp Ser Ile Tyr Ala Gly Ala Ser Gln Tyr Asp Leu Asp
  65     70     75
Tyr Val Phe Ala His Phe Ser Gly Ile Phe Leu Thr Ser Thr Val
  80     85     90
Tyr Phe Leu Ala Tyr Cys Ile Ala Met Lys Asn Ser Pro Lys Leu
  95    100    105
Tyr Pro Glu Ala Val Leu Pro Gly Phe Leu Ser Gly Val Leu Trp
  110    115    120
Ala Ile Ala Thr Cys Cys Trp Phe Ile Ala Asn His Ser Leu Ser
  125    130    135
Ala Val Val Ser Phe Pro Ile Ile Thr Ala Gly Pro Gly Phe Ile
  140    145    150
Ala Ala Met Trp Gly Ile Phe Met Phe Lys Glu Ile Lys Gly Leu
  155    160    165
Gln Asn Tyr Leu Leu Met Ile Leu Ala Phe Cys Ile Ile Leu Thr
  170    175    180
Gly Ala Leu Cys Thr Ala Phe Ser Lys Ile
  185    190

```

&lt;210&gt; 18

&lt;211&gt; 361

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

PF-0681 PCT

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 778800CD1

&lt;400&gt; 18

Met	Leu	Ala	Arg	Arg	Leu	Val	Trp	Ala	Leu	Ile	Ser	Glu	Ala	Thr	1	5	10	15
Lys	Ala	Gly	Ala	Ala	Ser	Met	Ile	His	Tyr	Met	Val	Leu	Ile	Ser	20	25	30	35
Ala	Arg	Leu	Val	Leu	Leu	Thr	Leu	Cys	Gly	Trp	Val	Leu	Cys	Trp	40	45	50	55
Thr	Leu	Val	Asn	Leu	Phe	Arg	Ser	His	Ser	Val	Leu	Asn	Leu	Leu	60	65	70	75
Phe	Leu	Gly	Tyr	Pro	Phe	Gly	Val	Tyr	Val	Pro	Leu	Cys	Cys	Phe	80	85	90	95
His	Gln	Asp	Ser	Arg	Ala	His	Leu	Leu	Leu	Thr	Asp	Tyr	Asn	Tyr	100	105	110	115
Val	Val	Gln	His	Glu	Ala	Val	Glu	Glu	Ser	Ala	Ser	Thr	Val	Gly	120	125	130	135
Gly	Leu	Ala	Lys	Ser	Lys	Asp	Phe	Leu	Ser	Leu	Leu	Leu	Glu	Ser	140	145	150	155
Leu	Lys	Glu	Gln	Phe	Asn	Asn	Ala	Thr	Pro	Ile	Pro	Thr	His	Ser	160	165	170	175
Cys	Pro	Leu	Ser	Pro	Asp	Leu	Ile	Arg	Asn	Glu	Val	Glu	Cys	Leu	180	185	190	195
Lys	Ala	Asp	Phe	Asn	His	Arg	Ile	Lys	Glu	Val	Leu	Phe	Asn	Ser	200	205	210	215
Leu	Phe	Ser	Ala	Tyr	Tyr	Val	Ala	Phe	Leu	Pro	Leu	Cys	Phe	Val	220	225	230	235
Lys	Ser	Thr	Gln	Tyr	Tyr	Asp	Met	Arg	Trp	Ser	Cys	Glu	His	Leu	240	245	250	255
Ile	Met	Val	Trp	Ile	Asn	Ala	Phe	Val	Met	Leu	Thr	Thr	Gln	Leu	260	265	270	275
Leu	Pro	Ser	Lys	Tyr	Cys	Asp	Leu	Leu	His	Lys	Ser	Ala	Ala	His	280	285	290	295
Leu	Gly	Lys	Trp	Gln	Lys	Leu	Glu	His	Gly	Ser	Tyr	Ser	Asn	Ala	300	305	310	315
Pro	Gln	His	Ile	Trp	Ser	Glu	Asn	Thr	Ile	Trp	Pro	Gln	Gly	Val	320	325	330	335
Leu	Val	Arg	His	Ser	Arg	Cys	Leu	Tyr	Arg	Ala	Met	Gly	Pro	Tyr	340	345	350	355
Asn	Val	Ala	Val	Pro	Ser	Asp	Val	Ser	His	Ala	Arg	Phe	Tyr	Phe	360			
Leu	Phe	His	Arg	Pro	Leu	Arg	Leu	Leu	Asn	Leu	Leu	Ile	Leu	Ile				
Glu	Gly	Ser	Val	Val	Phe	Tyr	Gln	Leu	Tyr	Ser	Leu	Leu	Arg	Ser				
Glu	Lys	Trp	Asn	His	Thr	Leu	Ser	Met	Ala	Leu	Ile	Leu	Phe	Cys				
Asn	Tyr	Tyr	Val	Leu	Phe	Lys	Leu	Leu	Arg	Asp	Arg	Ile	Val	Leu				
Gly	Arg	Ala	Tyr	Ser	Tyr	Pro	Leu	Asn	Ser	Tyr	Glu	Leu	Lys	Thr				

Asn

&lt;210&gt; 19

&lt;211&gt; 97

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

PF-0681 PCT

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1396995CD1

&lt;400&gt; 19

Met	Glu	Asp	His	Ile	Ile	Ser	Cys	Tyr	Leu	Lys	Trp	Pro	Val	Cys	
1				5					10					15	
Ala	Gln	Leu	Leu	Asn	Cys	Ser	Glu	Pro	Thr	Glu	Arg	Ala	Thr	Val	
				20					25					30	
Glu	Thr	Cys	Met	Val	Ser	Arg	Ile	His	Arg	His	Ala	Leu	Phe	Leu	
				35					40					45	
Cys	Leu	Leu	Pro	Gln	Cys	Tyr	Leu	Pro	Lys	Asp	Leu	Trp	Tyr	Ser	
				50					55					60	
Ser	Leu	Ser	Trp	Glu	Arg	Asn	Ser	Ile	Leu	Phe	Ser	Leu	Leu	Leu	
				65					70					75	
Leu	Tyr	Ser	Leu	Phe	Tyr	Ser	Tyr	Ala	Pro	Ile	Val	Ala	Lys	Gly	
				80					85					90	
Lys	Arg	Gly	Lys	Met	Ile	Val									
				95											

&lt;210&gt; 20

&lt;211&gt; 232

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1597730CD1

&lt;400&gt; 20

Met	Gly	Pro	Leu	Ala	Leu	Pro	Ala	Trp	Leu	Gln	Pro	Arg	Tyr	Arg	
1				5					10					15	
Lys	Asn	Ala	Tyr	Leu	Phe	Ile	Tyr	Tyr	Leu	Ile	Gln	Phe	Cys	Gly	
				20					25					30	
His	Ser	Trp	Ile	Phe	Thr	Asn	Met	Thr	Val	Arg	Phe	Phe	Ser	Phe	
				35					40					45	
Gly	Lys	Asp	Ser	Met	Val	Asp	Thr	Phe	Tyr	Ala	Ile	Gly	Leu	Val	
				50					55					60	
Met	Arg	Leu	Cys	Gln	Ser	Val	Ser	Leu	Leu	Glu	Leu	Leu	His	Ile	
				65					70					75	
Tyr	Val	Gly	Ile	Glu	Ser	Asn	His	Leu	Leu	Pro	Arg	Phe	Leu	Gln	
				80					85					90	
Leu	Thr	Glu	Arg	Ile	Ile	Ile	Leu	Phe	Val	Val	Ile	Thr	Ser	Gln	
				95					100					105	
Glu	Glu	Val	Gln	Glu	Lys	Tyr	Val	Val	Cys	Val	Leu	Phe	Val	Phe	
				110					115					120	
Trp	Asn	Leu	Leu	Asp	Met	Val	Arg	Tyr	Thr	Tyr	Ser	Met	Leu	Ser	
				125					130					135	
Val	Ile	Gly	Ile	Ser	Tyr	Ala	Val	Leu	Thr	Trp	Leu	Ser	Gln	Thr	
				140					145					150	
Leu	Trp	Met	Pro	Ile	Tyr	Pro	Leu	Cys	Val	Leu	Ala	Glu	Ala	Phe	
				155					160					165	
Ala	Ile	Tyr	Gln	Ser	Leu	Pro	Tyr	Phe	Glu	Ser	Phe	Gly	Thr	Tyr	
				170					175					180	
Ser	Thr	Lys	Leu	Pro	Phe	Asp	Leu	Ser	Ile	Tyr	Phe	Pro	Tyr	Val	
				185					190					195	
Leu	Lys	Ile	Tyr	Leu	Met	Met	Leu	Phe	Ile	Gly	Met	Tyr	Phe	Thr	
				200					205					210	
Tyr	Ser	His	Leu	Tyr	Ser	Glu	Arg	Arg	Asp	Ile	Leu	Gly	Ile	Phe	
				215					220					225	
Pro	Ile	Lys	Lys	Lys	Lys	Met									
				230											

&lt;210&gt; 21

PF-0681 PCT

&lt;211&gt; 271

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1629304CD1

&lt;400&gt; 21

Met	Thr	Ala	Trp	Ile	Leu	Leu	Pro	Val	Ser	Leu	Ser	Ala	Phe	Ser	1	5	10	15
Ile	Thr	Gly	Ile	Trp	Thr	Val	Tyr	Ala	Met	Ala	Val	Met	Asn	His	20	25	30	35
His	Val	Cys	Pro	Val	Glu	Asn	Trp	Ser	Tyr	Asn	Glu	Ser	Cys	Pro	40	45	50	55
Pro	Asp	Pro	Ala	Glu	Gln	Gly	Gly	Pro	Lys	Thr	Cys	Cys	Thr	Leu	60	65	70	75
Asp	Asp	Val	Pro	Leu	Ile	Ser	Lys	Cys	Gly	Ser	Tyr	Pro	Pro	Glu	80	85	90	95
Ser	Cys	Leu	Phe	Ser	Leu	Ile	Gly	Asn	Met	Gly	Ala	Phe	Met	Val	100	105	110	115
Ala	Leu	Ile	Cys	Leu	Leu	Arg	Tyr	Gly	Gln	Leu	Leu	Glu	Gln	Ser	120	125	130	135
Arg	His	Ser	Trp	Val	Asn	Thr	Thr	Ala	Leu	Ile	Thr	Gly	Cys	Thr	140	145	150	155
Asn	Ala	Ala	Gly	Leu	Leu	Val	Val	Gly	Asn	Phe	Gln	Val	Asp	His	160	165	170	175
Ala	Arg	Ser	Leu	His	Tyr	Val	Gly	Ala	Gly	Val	Ala	Phe	Pro	Ala	180	185	190	195
Gly	Leu	Leu	Phe	Val	Cys	Leu	His	Cys	Ala	Leu	Ser	Tyr	Gln	Gly	200	205	210	215
Ala	Thr	Ala	Pro	Leu	Asp	Leu	Ala	Val	Ala	Tyr	Leu	Arg	Ser	Val	220	225	230	235
Leu	Ala	Val	Ile	Ala	Phe	Ile	Thr	Leu	Val	Leu	Ser	Gly	Val	Phe	240	245	250	255
Phe	Val	His	Glu	Ser	Ser	Gln	Leu	Gln	His	Gly	Ala	Ala	Leu	Cys	260	265	270	
Glu	Trp	Val	Cys	Val	Ile	Asp	Ile	Leu	Ile	Phe	Tyr	Gly	Thr	Phe				
Ser	Tyr	Glu	Phe	Gly	Ala	Val	Ser	Ser	Asp	Thr	Leu	Val	Ala	Ala				
Leu	Gln	Pro	Thr	Pro	Gly	Arg	Ala	Cys	Lys	Ser	Ser	Gly	Ser	Ser				
Ser	Thr	Ser	Thr	His	Leu	Asn	Cys	Ala	Pro	Glu	Ser	Ile	Ala	Met				
Ile																		

&lt;210&gt; 22

&lt;211&gt; 267

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1989585CD1

&lt;400&gt; 22

Met	Ala	Pro	Ala	Ala	Asp	Arg	Glu	Gly	Tyr	Trp	Gly	Pro	Thr	Thr	1	5	10	15
Ser	Thr	Leu	Asp	Trp	Cys	Glu	Glu	Asn	Tyr	Ser	Val	Thr	Trp	Tyr	20	25	30	



PF-0681 PCT

```

Ile Ala Glu Phe Trp Asn Thr Val Ser Asn Leu Ile Met Ile Ile
      35      40      45
Pro Pro Met Phe Gly Ala Ile Gln Ser Val Arg Asp Gly Leu Glu
      50      55      60
Lys Arg Tyr Ile Ala Ser Tyr Leu Ala Leu Thr Val Val Gly Met
      65      70      75
Gly Ser Trp Cys Phe His Met Thr Leu Lys Tyr Glu Met Gln Leu
      80      85      90
Leu Asp Glu Leu Pro Met Ile Tyr Ser Cys Cys Ile Phe Val Tyr
      95     100     105
Cys Met Phe Glu Cys Phe Lys Ile Lys Asn Ser Val Asn Tyr His
     110     115     120
Leu Leu Phe Thr Leu Val Leu Phe Ser Leu Ile Val Thr Thr Val
     125     130     135
Tyr Leu Lys Val Lys Glu Pro Ile Phe His Gln Val Met Tyr Gly
     140     145     150
Met Leu Val Phe Thr Leu Val Leu Arg Ser Ile Tyr Ile Val Thr
     155     160     165
Trp Val Tyr Pro Trp Leu Arg Gly Leu Gly Tyr Thr Ser Leu Gly
     170     175     180
Ile Phe Leu Leu Gly Phe Leu Phe Trp Asn Ile Asp Asn Ile Phe
     185     190     195
Cys Glu Ser Leu Arg Asn Phe Arg Lys Lys Val Pro Pro Ile Ile
     200     205     210
Gly Ile Thr Thr Gln Phe His Ala Trp Trp His Ile Leu Thr Gly
     215     220     225
Leu Gly Ser Tyr Leu His Ile Leu Phe Ser Leu Tyr Thr Arg Thr
     230     235     240
Leu Tyr Leu Arg Tyr Arg Pro Lys Val Lys Phe Leu Phe Gly Ile
     245     250     255
Trp Pro Val Ile Leu Phe Glu Pro Leu Arg Lys His
     260     265

```

&lt;210&gt; 23

&lt;211&gt; 406

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2134729CD1

&lt;400&gt; 23

```

Met Asp Gln Ser Gly Met Glu Ile Pro Val Thr Leu Ile Ile Lys
  1      5      10      15
Ala Pro Asn Gln Lys Tyr Ser Asp Gln Thr Ile Ser Cys Phe Leu
     20      25      30
Asn Trp Thr Val Gly Lys Leu Lys Thr His Leu Ser Asn Val Tyr
     35      40      45
Pro Ser Lys Pro Leu Thr Lys Asp Gln Arg Leu Val Tyr Ser Gly
     50      55      60
Arg Leu Leu Pro Asp His Leu Gln Leu Lys Asp Ile Leu Arg Lys
     65      70      75
Gln Asp Glu Tyr His Met Val His Leu Val Cys Thr Ser Arg Thr
     80      85      90
Pro Pro Ser Ser Pro Lys Ser Ser Thr Asn Arg Glu Ser His Glu
     95     100     105
Ala Leu Ala Ser Ser Ser Asn Ser Ser Ser Asp His Ser Gly Ser
    110     115     120
Thr Thr Pro Ser Ser Gly Gln Glu Thr Leu Ser Leu Ala Val Gly
    125     130     135
Ser Ser Ser Glu Gly Leu Arg Gln Arg Thr Leu Pro Gln Ala Gln

```

PF-0681 PCT

Thr	Asp	Gln	Ala	140	Gln	Ser	His	Gln	Phe	145	Pro	Tyr	Val	Met	Gln	Gly	150
				155						160							165
Asn	Val	Asp	Asn	170	Gln	Phe	Pro	Gly	Gln	175	Ala	Ala	Pro	Pro	Gly	Phe	180
Pro	Val	Tyr	Pro	185	Ala	Phe	Ser	Pro	Leu	190	Gln	Met	Leu	Trp	Trp	Gln	195
Gln	Met	Tyr	Ala	200	His	Gln	Tyr	Tyr	Met	205	Gln	Tyr	Gln	Ala	Ala	Val	210
Ser	Ala	Gln	Ala	215	Thr	Ser	Asn	Val	Asn	220	Pro	Thr	Gln	Pro	Thr	Thr	225
Ser	Gln	Pro	Leu	230	Asn	Leu	Ala	His	Val	235	Pro	Gly	Glu	Glu	Pro	Pro	240
Pro	Ala	Pro	Asn	245	Leu	Val	Ala	Gln	Glu	250	Asn	Arg	Pro	Met	Asn	Glu	255
Asn	Val	Gln	Met	260	Asn	Ala	Gln	Gly	Gly	265	Pro	Val	Leu	Asn	Glu	Glu	270
Asp	Phe	Asn	Arg	275	Asp	Trp	Leu	Asp	Trp	280	Met	Tyr	Thr	Phe	Ser	Arg	285
Ala	Ala	Ile	Leu	290	Leu	Ser	Ile	Val	Tyr	295	Phe	Tyr	Ser	Ser	Phe	Ser	300
Arg	Phe	Ile	Met	305	Val	Met	Gly	Ala	Met	310	Leu	Leu	Val	Tyr	Leu	His	315
Gln	Ala	Gly	Trp	320	Phe	Pro	Phe	Arg	Gln	325	Glu	Gly	Gly	His	Gln	Gln	330
Ala	Pro	Asn	Asn	335	Asn	Ala	Glu	Val	Asn	340	Asn	Asp	Gly	Gln	Asn	Ala	345
Asn	Asn	Leu	Glu	350	Leu	Glu	Glu	Met	Glu	355	Arg	Leu	Met	Asp	Asp	Gly	360
Leu	Glu	Asp	Glu	365	Ser	Gly	Glu	Asp	Gly	370	Gly	Glu	Asp	Ala	Ser	Ala	375
Ile	Gln	Arg	Pro	380	Gly	Leu	Met	Ala	Ser	385	Ala	Trp	Ser	Phe	Ile	Thr	390
Thr	Phe	Phe	Thr	395	Ser	Leu	Ile	Pro	Glu	400	Gly	Pro	Pro	Gln	Val	Ala	405

Asn

&lt;210&gt; 24

&lt;211&gt; 318

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2299506CD1

&lt;400&gt; 24

Met	Leu	Ser	Glu	Ser	Ser	Phe	Leu	Lys	Gly	Val	Met	Leu	Gly	1	5	10	15
Ser	Ile	Phe	Cys	Ala	Leu	Ile	Thr	Met	Leu	Gly	His	Ile	Arg	Ile	20	25	30
Gly	His	Gly	Asn	Arg	Met	His	His	His	Glu	His	His	His	Leu	Gln	35	40	45
Ala	Pro	Asn	Lys	Glu	Asp	Ile	Leu	Lys	Ile	Ser	Glu	Asp	Glu	Arg	50	55	60
Met	Glu	Leu	Ser	Lys	Ser	Phe	Arg	Val	Tyr	Cys	Ile	Ile	Leu	Val	65	70	75
Lys	Pro	Lys	Asp	Val	Ser	Leu	Trp	Ala	Ala	Val	Lys	Glu	Thr	Trp	80	85	90
Thr	Lys	His	Cys	Asp	Lys	Ala	Glu	Phe	Phe	Ser	Ser	Glu	Asn	Val	95	100	105

PF-0681 PCT

Lys Val Phe Glu Ser Ile Asn Met Asp Thr Asn Asp Met Trp Leu  
 110 115 120  
 Met Met Arg Lys Ala Tyr Lys Tyr Ala Phe Asp Lys Tyr Arg Asp  
 125 130 135  
 Gln Tyr Asn Trp Phe Phe Leu Ala Arg Pro Thr Thr Phe Ala Ile  
 140 145 150  
 Ile Glu Asn Leu Lys Tyr Phe Leu Leu Lys Lys Asp Pro Ser Gln  
 155 160 165  
 Pro Phe Tyr Leu Gly His Thr Ile Lys Ser Gly Asp Leu Glu Tyr  
 170 175 180  
 Val Gly Met Glu Gly Gly Ile Val Leu Ser Val Glu Ser Met Lys  
 185 190 195  
 Arg Leu Asn Ser Leu Leu Asn Ile Pro Glu Lys Cys Pro Glu Gln  
 200 205 210  
 Gly Gly Met Ile Trp Lys Ile Ser Glu Asp Lys Gln Leu Ala Val  
 215 220 225  
 Cys Leu Lys Tyr Ala Gly Val Phe Ala Glu Asn Ala Glu Asp Ala  
 230 235 240  
 Asp Gly Lys Asp Val Phe Asn Thr Lys Ser Val Gly Leu Ser Ile  
 245 250 255  
 Lys Glu Ala Met Thr Tyr His Pro Asn Gln Val Val Glu Gly Cys  
 260 265 270  
 Cys Ser Asp Met Ala Val Thr Phe Asn Gly Leu Thr Pro Asn Gln  
 275 280 285  
 Met His Val Met Met Tyr Gly Val Tyr Arg Leu Arg Ala Phe Gly  
 290 295 300  
 His Ile Phe Asn Asp Ala Leu Val Phe Leu Pro Pro Asn Gly Ser  
 305 310 315  
 Asp Asn Asp

&lt;210&gt; 25

&lt;211&gt; 326

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2506558CD1

&lt;400&gt; 25

Met Glu Phe Leu Trp Asp Phe Leu Asn His Gln Glu Gly Pro Arg  
 1 5 10 15  
 Ile Arg Asp His Leu Ser His Gly Glu Ile Asn Leu His Glu Phe  
 20 25 30  
 Ser Lys Glu Thr Thr Asn Gln Leu Leu Ala Phe Ser Leu Val Leu  
 35 40 45  
 Leu Leu Arg Phe Val Asp Asp Cys Leu Leu Ser Val Phe Lys Glu  
 50 55 60  
 Lys Ser Ala Val Glu Leu Leu Ile Ser Leu Ala Glu Gly Tyr Ser  
 65 70 75  
 Ser Arg Cys His Pro Val Phe Gln Leu Lys Lys Gln Val Leu Ser  
 80 85 90  
 Cys Glu Glu Ser Ile Arg Val Trp Ala Leu Leu Pro Phe Pro Glu  
 95 100 105  
 Glu Leu Thr Arg Gln Ala Val Arg Leu Glu Asp Asn Ser Glu Thr  
 110 115 120  
 Asn Ala Cys His Ser Leu Ile Thr Lys Met Thr Asp Glu Leu Tyr  
 125 130 135  
 His His Met Pro Glu Asn Arg Cys Val Leu Lys Asp Leu Asp Arg  
 140 145 150  
 Leu Pro Thr Glu Thr Trp Pro Gln Leu Leu Arg Glu Leu Cys Ser

PF-0681 PCT

Thr	Pro	Val	Pro	Thr	Leu	Phe	Cys	Pro	Arg	Ile	Val	Leu	Glu	Val	155	160	165
Leu	Val	Val	Leu	Arg	Ser	Ile	Gly	Glu	Gln	Cys	Arg	Arg	Val	Ser	170	175	180
Ser	Gln	Val	Thr	Val	Ala	Ser	Glu	Leu	Arg	His	Arg	Gln	Trp	Val	185	190	195
Glu	Arg	Thr	Leu	Arg	Ser	Arg	Gln	Arg	Gln	Asn	Tyr	Leu	Arg	Met	200	205	210
Trp	Ser	Ser	Ile	Arg	Leu	Leu	Ser	Pro	Val	Leu	Ser	Leu	Ile	Leu	215	220	225
Leu	Leu	Ile	Ala	Leu	Glu	Leu	Val	Asn	Ile	His	Ala	Val	Cys	Gly	230	235	240
Lys	Asn	Ala	His	Glu	Tyr	Gln	Gln	Tyr	Leu	Lys	Phe	Val	Lys	Ser	245	250	255
Ile	Leu	Gln	Tyr	Thr	Glu	Asn	Leu	Val	Ala	Tyr	Thr	Ser	Tyr	Glu	260	265	270
Lys	Asn	Lys	Trp	Asn	Glu	Thr	Ile	Asn	Leu	Thr	His	Thr	Ala	Leu	275	280	285
Leu	Lys	Met	Trp	Thr	Phe	Ser	Glu	Lys	Lys	Gln	Met	Leu	Ile	His	290	295	300
Leu	Ala	Lys	Lys	Ser	Thr	Ser	Lys	Val	Leu	Leu					305	310	315
															320	325	

&lt;210&gt; 26

&lt;211&gt; 247

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2546025CD1

&lt;400&gt; 26

Met	Leu	Arg	Phe	Ile	Gln	Lys	Phe	Ser	Gln	Ala	Ser	Ser	Lys	Ile			
1				5					10					15			
Leu	Lys	Tyr	Ser	Phe	Pro	Val	Gly	Leu	Arg	Thr	Ser	Arg	Thr	Asp			
				20					25					30			
Ile	Leu	Ser	Leu	Lys	Met	Ser	Leu	Gln	Gln	Asn	Phe	Ser	Pro	Cys			
				35					40					45			
Pro	Arg	Pro	Trp	Leu	Ser	Ser	Ser	Phe	Pro	Ala	Tyr	Met	Ser	Lys			
				50					55					60			
Thr	Gln	Cys	Tyr	His	Thr	Ser	Pro	Cys	Ser	Phe	Lys	Lys	Gln	Gln			
				65					70					75			
Lys	Gln	Ala	Leu	Leu	Ala	Arg	Pro	Ser	Ser	Thr	Ile	Thr	Tyr	Leu			
				80					85					90			
Thr	Asp	Ser	Pro	Lys	Pro	Ala	Leu	Cys	Val	Thr	Leu	Ala	Gly	Leu			
				95					100					105			
Ile	Pro	Phe	Val	Ala	Pro	Pro	Leu	Val	Met	Leu	Met	Thr	Lys	Thr			
				110					115					120			
Tyr	Ile	Pro	Ile	Leu	Ala	Phe	Thr	Gln	Met	Ala	Tyr	Gly	Ala	Ser			
				125					130					135			
Phe	Leu	Ser	Phe	Leu	Gly	Gly	Ile	Arg	Trp	Gly	Phe	Ala	Leu	Pro			
				140					145					150			
Glu	Gly	Ser	Pro	Ala	Lys	Pro	Asp	Tyr	Leu	Asn	Leu	Ala	Ser	Ser			
				155					160					165			
Ala	Ala	Pro	Leu	Phe	Phe	Ser	Trp	Phe	Ala	Phe	Leu	Ile	Ser	Glu			
				170					175					180			
Arg	Leu	Ser	Glu	Ala	Ile	Val	Thr	Val	Ile	Met	Gly	Met	Gly	Val			
				185					190					195			
Ala	Phe	His	Leu	Glu	Leu	Phe	Leu	Leu	Pro	His	Tyr	Pro	Asn	Trp			
				200					205					210			

PF-0681 PCT

Phe	Lys	Ala	Leu	Arg	Ile	Val	Val	Thr	Leu	Leu	Ala	Thr	Phe	Ser
			215						220					225
Phe	Ile	Ile	Thr	Leu	Val	Val	Lys	Ser	Ser	Phe	Pro	Glu	Lys	Gly
			230						235					240
His	Lys	Arg	Pro	Gly	Gln	Val								
				245										

&lt;210&gt; 27

&lt;211&gt; 278

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3145660CD1

&lt;400&gt; 27

Met	Ala	Tyr	Phe	Tyr	Met	Cys	Asp	Arg	Ala	Asn	Leu	Phe	Met	Lys
1				5					10					15
Glu	Asn	Lys	Phe	Tyr	Thr	His	Ser	Ser	Phe	Phe	Ile	Pro	Ile	Ile
				20					25					30
Tyr	Ile	Leu	Val	Leu	Gly	Val	Phe	Tyr	Asn	Glu	Asn	Thr	Lys	Glu
				35					40					45
Thr	Lys	Val	Leu	Asn	Arg	Glu	Gln	Thr	Asp	Glu	Trp	Lys	Gly	Trp
				50					55					60
Met	Gln	Leu	Val	Ile	Leu	Ile	Tyr	His	Ile	Ser	Gly	Ala	Ser	Thr
				65					70					75
Phe	Leu	Pro	Val	Tyr	Met	His	Ile	Arg	Val	Leu	Val	Ala	Ala	Tyr
				80					85					90
Leu	Phe	Gln	Thr	Gly	Tyr	Gly	His	Phe	Ser	Tyr	Phe	Trp	Ile	Lys
				95					100					105
Gly	Asp	Phe	Gly	Ile	Tyr	Arg	Val	Cys	Gln	Val	Leu	Phe	Arg	Leu
				110					115					120
Asn	Phe	Leu	Val	Val	Val	Leu	Cys	Ile	Val	Met	Asp	Arg	Pro	Tyr
				125					130					135
Gln	Phe	Tyr	Tyr	Phe	Val	Pro	Leu	Val	Thr	Val	Trp	Phe	Met	Val
				140					145					150
Ile	Tyr	Val	Thr	Leu	Ala	Leu	Trp	Pro	Gln	Ile	Ile	Gln	Lys	Lys
				155					160					165
Ala	Asn	Gly	Asn	Cys	Phe	Trp	His	Phe	Gly	Leu	Leu	Leu	Lys	Leu
				170					175					180
Gly	Phe	Leu	Leu	Phe	Ile	Cys	Phe	Phe	Leu	Ala	Tyr	Ser	Gln	Gly
				185					190					195
Ala	Phe	Glu	Lys	Ile	Phe	Ser	Leu	Trp	Pro	Leu	Ser	Lys	Cys	Phe
				200					205					210
Glu	Leu	Lys	Gly	Asn	Val	Tyr	Glu	Trp	Trp	Phe	Arg	Trp	Arg	Leu
				215					220					225
Asp	Arg	Tyr	Val	Val	Phe	His	Gly	Met	Leu	Phe	Ala	Phe	Ile	Tyr
				230					235					240
Leu	Ala	Leu	Gln	Lys	Arg	Gln	Ile	Leu	Ser	Glu	Gly	Lys	Gly	Glu
				245					250					255
Pro	Leu	Phe	Ser	Asn	Lys	Ile	Ser	Asn	Phe	Leu	Leu	Val	Ile	Ser
				260					265					270
Val	Val	Ser	Phe	Leu	Gly	Lys	Phe							
				275										

&lt;210&gt; 28

&lt;211&gt; 320

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

PF-0681 PCT

&lt;223&gt; Incyte ID No: 4901066CD1

&lt;400&gt; 28

Met	Thr	Leu	Trp	Asn	Gly	Val	Leu	Pro	Phe	Tyr	Pro	Gln	Pro	Arg	1	5	10	15
His	Ala	Ala	Gly	Phe	Ser	Val	Pro	Leu	Leu	Ile	Val	Ile	Leu	Val	20	25	30	35
Phe	Leu	Ala	Leu	Ala	Ala	Ser	Phe	Leu	Leu	Ile	Leu	Pro	Gly	Ile	40	45	50	55
Arg	Gly	His	Ser	Arg	Trp	Phe	Trp	Leu	Val	Arg	Val	Leu	Leu	Ser	60	65	70	75
Leu	Phe	Ile	Gly	Ala	Glu	Ile	Val	Ala	Val	His	Phe	Ser	Ala	Glu	80	85	90	95
Trp	Phe	Val	Gly	Thr	Val	Asn	Thr	Asn	Thr	Ser	Tyr	Lys	Ala	Phe	100	105	110	115
Ser	Ala	Ala	Arg	Val	Thr	Ala	Arg	Val	Gly	Leu	Leu	Val	Gly	Leu	120	125	130	135
Glu	Gly	Ile	Asn	Ile	Thr	Leu	Thr	Gly	Thr	Pro	Val	His	Gln	Leu	140	145	150	155
Asn	Glu	Thr	Ile	Asp	Tyr	Asn	Glu	Gln	Phe	Thr	Trp	Arg	Leu	Lys	160	165	170	175
Glu	Asn	Tyr	Ala	Ala	Glu	Tyr	Ala	Asn	Ala	Leu	Glu	Lys	Gly	Leu	180	185	190	195
Pro	Asp	Pro	Val	Leu	Tyr	Leu	Ala	Glu	Lys	Phe	Thr	Pro	Ser	Ser	200	205	210	215
Pro	Cys	Gly	Leu	Tyr	His	Gln	Tyr	His	Leu	Ala	Gly	His	Tyr	Ala	220	225	230	235
Ser	Ala	Thr	Leu	Trp	Val	Ala	Phe	Cys	Phe	Trp	Leu	Leu	Ser	Asn	240	245	250	255
Val	Leu	Leu	Ser	Thr	Pro	Ala	Pro	Leu	Tyr	Gly	Gly	Leu	Ala	Leu	260	265	270	275
Leu	Thr	Thr	Gly	Ala	Phe	Ala	Leu	Phe	Gly	Val	Phe	Ala	Leu	Ala	280	285	290	295
Ser	Ile	Ser	Ser	Val	Pro	Leu	Cys	Pro	Leu	Arg	Leu	Gly	Ser	Ser	300	305	310	315
Ala	Leu	Thr	Thr	Gln	Tyr	Gly	Ala	Ala	Phe	Trp	Val	Thr	Leu	Ala	320			
Thr	Gly	Val	Leu	Cys	Leu	Phe	Leu	Gly	Gly	Ala	Val	Val	Ser	Leu				
Gln	Tyr	Val	Arg	Pro	Ser	Ala	Leu	Arg	Thr	Leu	Leu	Asp	Gln	Ser				
Ala	Lys	Asp	Cys	Ser	Gln	Glu	Arg	Gly	Gly	Ser	Pro	Leu	Ile	Leu				
Gly	Asp	Pro	Leu	His	Lys	Gln	Ala	Ala	Leu	Pro	Asp	Leu	Lys	Cys				
Ile	Thr	Thr	Asn	Leu														

&lt;210&gt; 29

&lt;211&gt; 360

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 5031174CD1

&lt;400&gt; 29

Met	Ser	Gln	Gly	Arg	Gly	Lys	Tyr	Asp	Phe	Tyr	Ile	Gly	Leu	Gly	1	5	10	15
Leu	Ala	Met	Ser	Ser	Ser	Ile	Phe	Ile	Gly	Gly	Ser	Phe	Ile	Leu	20	25	30	35

PF-0681 PCT

Lys	Lys	Lys	Gly	Leu	Leu	Arg	Leu	Ala	Arg	Lys	Gly	Ser	Met	Arg	35	40	45
Ala	Gly	Gln	Gly	Gly	His	Ala	Tyr	Leu	Lys	Glu	Trp	Leu	Trp	Trp	50	55	60
Ala	Gly	Leu	Leu	Ser	Met	Gly	Ala	Gly	Glu	Val	Ala	Asn	Phe	Ala	65	70	75
Ala	Tyr	Ala	Phe	Ala	Pro	Ala	Thr	Leu	Val	Thr	Pro	Leu	Gly	Ala	80	85	90
Leu	Ser	Val	Leu	Val	Ser	Ala	Ile	Leu	Ser	Ser	Tyr	Phe	Leu	Asn	95	100	105
Glu	Arg	Leu	Asn	Leu	His	Gly	Lys	Ile	Gly	Cys	Leu	Leu	Ser	Ile	110	115	120
Leu	Gly	Ser	Thr	Val	Met	Val	Ile	His	Ala	Pro	Lys	Glu	Glu	Glu	125	130	135
Ile	Glu	Thr	Leu	Asn	Glu	Met	Ser	His	Lys	Leu	Gly	Asp	Pro	Gly	140	145	150
Phe	Val	Val	Phe	Ala	Thr	Leu	Val	Val	Ile	Val	Ala	Leu	Ile	Leu	155	160	165
Ile	Phe	Val	Val	Gly	Pro	Arg	His	Gly	Gln	Thr	Asn	Ile	Leu	Val	170	175	180
Tyr	Ile	Thr	Ile	Cys	Ser	Val	Ile	Gly	Ala	Phe	Ser	Val	Ser	Cys	185	190	195
Val	Lys	Gly	Leu	Gly	Ile	Ala	Ile	Lys	Glu	Leu	Phe	Ala	Gly	Lys	200	205	210
Pro	Val	Leu	Arg	His	Pro	Leu	Ala	Trp	Ile	Leu	Leu	Leu	Ser	Leu	215	220	225
Ile	Val	Cys	Val	Ser	Thr	Gln	Ile	Asn	Tyr	Leu	Asn	Arg	Ala	Leu	230	235	240
Asp	Ile	Phe	Asn	Thr	Ser	Ile	Val	Thr	Pro	Ile	Tyr	Tyr	Val	Phe	245	250	255
Phe	Thr	Thr	Ser	Val	Leu	Thr	Cys	Ser	Ala	Ile	Leu	Phe	Lys	Glu	260	265	270
Trp	Gln	Asp	Met	Pro	Val	Asp	Asp	Val	Ile	Gly	Thr	Leu	Ser	Gly	275	280	285
Phe	Phe	Thr	Ile	Ile	Val	Gly	Ile	Phe	Leu	Leu	His	Ala	Phe	Lys	290	295	300
Asp	Val	Ser	Phe	Ser	Leu	Ala	Ser	Leu	Pro	Val	Ser	Phe	Arg	Lys	305	310	315
Asp	Glu	Lys	Ala	Met	Asn	Gly	Asn	Leu	Ser	Asn	Met	Tyr	Glu	Val	320	325	330
Leu	Asn	Asn	Asn	Glu	Glu	Ser	Leu	Thr	Cys	Gly	Ile	Glu	Gln	His	335	340	345
Thr	Gly	Glu	Asn	Val	Ser	Arg	Arg	Asn	Gly	Asn	Leu	Thr	Ala	Phe	350	355	360

&lt;210&gt; 30

&lt;211&gt; 1716

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1316927CB1

&lt;400&gt; 30

```

ccccgactct ggctgggctg aggagggggc cccacaggcc tgggtggaag cccgatccctg 60
agcaatcccc ctgtgcagga ttgaacatga ttgtctctct tttctcctt tttctgactt 120
taaggcaact ttgtccctaa caaaatgtgg aagggttagg tcaagaggaa tgcattctgtg 180
gaaacagttg ataataaaac gtctgaggat gtaaccatgg cagcagcttc tcctgtcaca 240
ttgaccaag ggacttcggc agcccacctc aactctatgg aagtcacaac agaggacaca 300
agcaggacag atgtgagtga accagcaact tcaggagggtg cagctgatgg tgtgacctcc 360

```

PF-0681 PCT

```

attgctccca cggctgtggc ctccagtacg actgcggcct ccattacgac tgcggcctcc 420
agtatgactg tggcctccag tgctcccacg actgcagcct ccagtacaac tgtggcctcc 480
attgctccca cgactgcagc ctccagtatg actgcggcct ccagcactcc catgacactt 540
gcactccccg cgcccacgtc cacttcaca gggcggaacc cgtccactac cgccactggg 600
catccatctc tcagcacagc cctcgacaaa gtgccaaaaga gcagcgcgtt gccaaagaaca 660
gcaaccttgg ccacattggc cacacgtgct cagactgtag cgaccacagc aaacacaagc 720
agccccatga gcaactgtcc aagtccttcc aagcacatgc ccagtgcacac cgcggaagc 780
cctgtacccc ctatgcgtcc ccaagcaca ggtcccatta gccaggtgtc agtggaccag 840
cctgtgggta acacaacaaa taaatccaca cccatgccct caaacacaac cccagagccc 900
gccccacccc ccacagtggg gaccaccacc aaggcacaag ccaggagacc aactgccagc 960
ccagtgccag tacctcacac cagcccaatc cctgagatgg aggccatgtc cccacgcaca 1020
cagccaagcc ccatgccata taccagagg gccgctggg caggcacatc ccaggcaccg 1080
gagcaggtag agactgaagc cacaccaggt actgattcca ctgggccaac acccaggagc 1140
tcagggggca ctaagatgcc agccacggac tctgtccagc ccagcaccca aggccagtac 1200
atggtgggtc ccactgagcc cctcacccag gccgtggtag acaaaactct ccttctgggtg 1260
gtgctgttac tcggggtgac ccttttcatc acagtcttgg ttttgtttgc cctgcaggcc 1320
tatgagagct acaagaagaa ggactacacc caggtggact acttaatcaa cgggatgtat 1380
gcggaactcag aaatgtgagg gggcgggggg cctggcgagg ggcctggccc cttcctcgtc 1440
ctttcctttt gcctttgaga ccaaaccaag tgcttccaaa ttcttttggg gcaattgagg 1500
agatatgcca gatgcttaaa cacatttaat tgctgtcaga ttaattccat gatcactaaa 1560
gagttgtctg ttttttcata tttatttttg taaatgattc tgtgcccagg agcagctggg 1620
ggttccacct cagggtgggg cgggcaggac cccgtctccc caggtgtcgg agcctgacct 1680
gaattaaagt actgactgct cgccaaaaaa aaaaaa 1716

```

&lt;210&gt; 31

&lt;211&gt; 1133

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 1094

&lt;223&gt; a or g or c or t, unknown, or other

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1354891CB1

&lt;400&gt; 31

```

caggtaggca agtgagcact gcagccaatg ccactgtatc ccttgacagg attctggatg 60
gtgaaatgga ccccaaagg atgtgttcct cactgctgtg ggtactgaat gccaaagaat 120
ttctgggctg gtgttgggat tgtgtgcagg ttcaggaggc aggaactcct cccacctgct 180
tctgcttctt gtctgcctt actggctctt tctccaactc tccatcacc agtctttctc 240
tttcaagtac ctgctattga ggagaacttg ctatagatgaca agcatttgct gaagccatgg 300
gatgctaaga agctatcctc atcctcttcc cgacctcggt cctgtgaagt ccttggaaat 360
aacatctttc catctcctga ccagcctgcc aatgtgcctg tctctccacc tgccatgaac 420
acgggggggct ccctacctga cctcaccaac ctgcactttc cccaccact gccaccccc 480
ctggaccctg aagagacagc ctaccctagc ctgagtgggg gcaacagtac ctccaatttg 540
acccacacca tgactcacct gggcatcagc aggggcatgg gcctgggccc aggtatgat 600
gcaccaggac ttcattcacc tctcagccac ccctcctgct agtctctcct aagcaatccc 660
aacctccagg cttccctgag cagtctcag cccagcttc agggctccca cagccacccc 720
tctctgcctg cctctcctt ggcccgccat gtactgccc ccacctcctt gggccacccc 780
tactcagtg ctcggctct cctctctcc tcttcgtcct cctccacttc atctcctgtt 840
ttggcgccc cctcttacc tgcttctacc cctggggcct ccccccacca ccgctgtgtg 900
cccctcagcc ccctgagttt gctcggggc ccagccgacg ccagaaggte ccaacaggag 960
ctgcccacac agttttcgcc aacaatgtca cccacttgt cttccatcac tcaggcgctc 1020
ccctggatac cagtaaaactg tccactgacc agcgggttaac cccataccca tacagctccc 1080
caagtctggt tctngctacc caggcccaca ccccaaagtc tctacagcag cca 1133

```

&lt;210&gt; 32

&lt;211&gt; 3324



PF-0681 PCT

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1415487CB1

&lt;400&gt; 32

```
tagatgacaa tagtgtattt atgctgtacc ttttctatgt ttagatgtgt gtagatgcac 60
aaataacttgc cgtgggtgtta agattgcctg cagtattcac accataacat gcactgcagg 120
ctttagacct aggagcagtc gactctagcg tccatcctgg gtgtgtattg ggctgtacca 180
tctgggttgg tgtaagtgcg cgctacaatg ttcctatggt cataaaatca cctaacgaca 240
catttctaac gacacgtccc tgtcggttaag caatgcatga ctgtactaac ctctcctgca 300
gttgctcaga tggggcgagc ctactgtttg tgcaggggcc ctggctgcta ggctggaagc 360
tcatacgtct cagtgcacatc tcacctccca ggcatgggtg atgcccacaa acgtttggct 420
ggagcaccct gaagtctctg ctaccatcca agctgggttg gagacgtgaa ccagaccaat 480
gaagggtctg ctgtacaatg attctctcgt cctattttat taacttcac taccttgcca 540
agagcacaaa aaaaacccatg ctaactttaa ctttgggtat tgcaattaca ttcctccttg 600
tttgtagtgg aacatttttt ccatatagct ccaatcctgc taatccgaag ccaaagagag 660
tgtttcttca gcatatgact agaaccattc atgacttgga aggaaatgca gttaaacggg 720
actctggaat atggatcaat ggggttgatt atactggaat ttctcacata acccctcaca 780
ttcctgagat caatgatagt atccgagctc actgtgagga gaatgcacct ctttgtgggt 840
ttccttggtg tcttccagtg cactttctga tcaggaaaaa ctgggtatctt cctgccccag 900
aagtttctcc aagaaatcct cctcatttcc gactcatatc caaagaacag acaccttggg 960
attctataaa attgactttt gaagcaacag gaccaagcca tatgtccttc tatgttcgag 1020
cccacaaagg gtcaaacactt tctcagtggg ctcttggcaa tggcacccca gtcacaagta 1080
aaggaggaga ctactttgtc ttttactccc atggactcca ggctctgca tggcagttct 1140
ggatagaaat gcaggtttca gaagaacatc ctgaaggaaat ggtcaccgtg gccattgctg 1200
cccactatct gtctggggaa gacaagagat cccctcaact ggatgctctg aagggaaaagt 1260
tcccagattg gacatttccc tctgcctggg tgtgcacctc cgatctcttt gtattttaat 1320
cttgtggatg agctctaagt acatgcccag tggatactcc atgtgacatg gtttctccct 1380
atgttacgtg gatgtttgta acgtaagtca atgaatttta atgatcatat gttcaaagag 1440
ctttctgggt taacgctttt cagggccaaag cactataagg gtttagctgt ggcgcagtg 1500
tgcatggcct gttgacactt gaaaatgcca gtcttttggc acttcagcac atgtgggtac 1560
tgccactaca cacacgtcat tttatatgac cttaaggaca aagccaacaa tccacttcaa 1620
tagctgcccc tttaggatca agaaagatgt cactgtcag agcattgtta atgagacaaa 1680
agttgtttcc aatttaagcc ccaaaacccat ttgttgtatt agtggatggg gggtaaaata 1740
tcattcactg aggtaatgat tccccttgag aatataactc tgtgtaggtc actggaaagt 1800
gattgccata gggctgggag agaagcattg cactcttgag gctgtagcct gtgtcaagct 1860
gtttcttcag gcagcctctc aaatgtgctt tgtctctctg tgcctgaggcc tggacctgt 1920
gctgagctgg tgactcactg tccctgacaag gatgcactgc tgtgctgctt 1980
tccctgaggtg gttttctatg cctgttttcc tctgaaacat gtctgttacc cctctccatc 2040
ttaccaagtt gaaaagggga atatttggtc acatacccct ctgggttttcg taggttcttt 2100
tgggttcagaa tattgtttgt gccagtacat gaccttaact tecttctcct gagcactgag 2160
ctgccatctg ggctattctg gggtagaagg aaggctggga gtgggtggga ttttataaat 2220
atttattctc ttttctttgt ttcataaggag tcttgtgtta tacaaggtta gtccttcacg 2280
gtataatctt actgatgcac tgggcctatc tttttgtttt ccagccagtt gaatagatta 2340
gtttttctca gtaacttact atccagcaga ctggctttcc tgagacttga ggttgtggct 2400
tatactggaa tgagaccact gtacgtgtag gtgggttcaga tccctgcgtaa tggcagcatg 2460
aggacttaaa aggtgggttt cattttgaag atggctatgt agcttgtaag gtgtatcaca 2520
gcagtacctc tcatggcttt ttgggtccag cagtgggggc attgggtgaga tcaatggtaa 2580
actgtgcaag ctttcttttt atcattagga aatgtgaaac gttggacaaa ttttgagttt 2640
taacaaggac aaaaagttga aagaaaaggc acagttaaca aaaaaggggtg gctagattta 2700
tcttgggtga tggaggaaat gagagaggaa tgctcttgaa aggtgggtctg tggatctgtc 2760
tgaatagaaa gagcacagta agtatgcatt gccggagaaa acgtccttga agctgcttct 2820
ctcatgtgta tgatgtgctt tttaaatcat gcccctcggt gctgcctaa tctgtgactc 2880
cctaaaaact aactgggccc atgtagatgg ggctgcaacc agagctgaat aacatgttag 2940
gctcacacat gcatcagcac tgcacactgg aatcattgct cttcctggac tttgtagaaa 3000
tcagtctcaa gtgcttcaag agtctggctc ctgctacttt tatctgtcag gtacacata 3060
aggtttgtag ggtttatatt ttgtatagaa tcacagttgt ggagaaaaag taataatttc 3120
tcaatgaatt ttaaaaatgg gcctattttc tatccccgtg gttcatctga tataatttag 3180
```

PF-0681 PCT

```

gttccctgtg aattcccccc ctctatggga aggatgcctt tactctttat cagtaataaa 3240
ttatgactgt tttcatattg ccttaggggt atttccccgg gtgaaaccat tggcttttgt 3300
gttggttttc cttagcataa gaag                                     3324

```

<210> 33  
 <211> 1023  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1693184CB1

```

<400> 33
tggcggaag tggcgagcg cgacggtgcg gggcgcgcg gcgacgggccc agcgggatgg 60
gagcagggtc ccccgggcgc ccgcagcctc gggagccgcg ctgaagccgg gggcgggacc 120
agcgcgagc cgacatgtgt ctgcgcctcg gaggcctgag tgtgggcgac ttccggaagg 180
tgctgatgaa gacaggcctg gtgctgggtg tgctgggcca tgtgagcttc atcacagctg 240
ccctgttcca tggcacagtg ctgcgctacg tgggcacccc tcaagatgcg gtggctctgc 300
agtactgctg ggtcaacatc ctctctgtca ctccgccat cgtgggtcatc acttcaggca 360
tcgcagccat cgtgttgtca cgctacctcc ctageacccc cctgcgctgg acagtgttta 420
gctcgagcgt ggcctgtgct ctctttctc tgacctgtgc cctcggcctc ttggcctcca 480
tcgccatgac ctttgccacc caggggcaagg cactgctggc tgcctgcact tttgggagct 540
ctgaactact ggccttcgca cctgactgtc ccttcgaccc cacacgcatt tatagctcca 600
gcctgtgect ctggggcatc gccctagtgc tctgcgtggc ggagaacgtg tttgctgtac 660
gctgtgctca gctcaccac cagctgctgg agctgaggcc ctggtggggg aaaagcagcc 720
accacatgat gcgggagaac ccagagctgg tggaggccg tgacctgctg agctgcacca 780
gctctgagcc tctgaccctc tgagagatga tgcctgccc agggccgatg gccactagga 840
ccctgcaagc aactctgctc tgtgaccagg ccaggattcc tggagctggc ctgagagggc 900
tcaatggacc ctcggggacc caagtggggc ttcaaccct ctccccacc acccagccca 960
ctgcactgaa atgagacttt attctgaaat tattaataaag aacagagatg ctcaaaaaaa 1020
aaa                                             1023

```

<210> 34  
 <211> 2091  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1695049CB1

```

<400> 34
tgaagccgca gccgcccggc taggccccgg gcggtcttag cccagggcgg cccgcggggc 60
gctgggcctg gctccccgct ccggtttccg ggcggcgcg tggccgctca ccatgcccgg 120
caagcaccag catttccagg aacctgaggt cggctgtgctc gggaaatact tcctgtttgg 180
cttcaacatt gtcttctggg tgctgggagc cctgttctct gctatcggcc tctgggcctg 240
gggtgagaag ggcgttctct cgaacatctc agcgtgaca gatctgggag gccttgaccc 300
cgtgtggctg tttgtggtag ttggaggcgt catgtcgggtg ctgggctttg ctggctgcat 360
tggggccctc cgggagaaca ccttctgtct caagttttcc tccgtgttcc tcggtctcat 420
cttcttctct gagctggcaa cagggatcct ggcctttgtc ttcaaggact ggattcgaga 480
ccagctcaac ctcttcatca acaacaacgt caaggcctac cgggacgaca ttgacctcca 540
gaacctcatt gactttgtct aggaatactg gtcttgctgc ggagcccag gcccacatga 600
ctggaacctc aatatctact tcaactgcac tgacttgaac cccagccggg agcgtgcgg 660
ggtgcccttc tcctgctgct tcagggaccc tgaggaggat gtcctcaaca cccagtgtgg 720
ctacgacgtc cggctcaaac tggagctgga gcagcagggc ttcattccaca ccaaaggctg 780
cgtgggccag tttgagaagt ggctgcagga caacctgatt gtggtggcgg gagtcttcat 840
gggcatcgcc ctctccaga tcttggcat ctgctggcc cagaacctcg tgagtacat 900
caaggcagtg aaagccaact ggagcaaatg gaatgatgac tttgaaaacc actggcttac 960
gcccaccatt tccgaggtcc tgtccacggc ggggcctcag cagaactctc tgactggggc 1020
ccctggcccc gcccaccca gccgacatgt tttctttggc ctgggtggtt tataacctga 1080

```

PF-0681 PCT

```
gccaaccttt aaaaattggt agatttcaca taaaagtcca gatccacagc ttctcttgaa 1140
gaatgaccac ctggctacgc cggctcttcg gtggcaacac tacctgggac actgacctcc 1200
cagtcaccaa gggccccagc tggcccgttc tactcaccta agtgccgect gagccttgta 1260
cactaggagc tggcctccca cctctgcagg gttatttcct gcacctcgag gccgctgcgg 1320
gccaatctgg agtgaaacac ggggacctga aggatggaga ggctggaccc cgctttgaag 1380
aggggtgcagc ctgggaaggg cggccttgct ggggactgcg gtgggagtag agtgcccagg 1440
agagggtctg aggggtggga tgggggtcag gacaattttg caaaagaagt agctggaagc 1500
catgggactg gcgggagcct gtttggggga tctggatggt tgactcctag gagtcaagtt 1560
cagcatcttc ccgctggctg cagagctgcc tgatgggcac tagagggcac gccagcccca 1620
cactccctgg gtctggcttc ctcccgcaac ctactctag tagagcctgt gcctgcctac 1680
tagcgtctct gggttcggag agtttgggaa ttctcagag ccaactggct caggcttggg 1740
aaggctggct gctgccctca gctccgcctc atcagctatg tgaaggggtg tgtgtggagt 1800
gaccttgccg cccctccctc gggctggctc agagatctca aactccgatg ccctggggc 1860
cacgtatggt ttgtaaatgg atgaaacagg cccttgagtt gggagcctgc ttacttttga 1920
ctttccact gttgctggag acaaagacat cgtgatgaga gaaagttcgc acaatctagt 1980
cggtaacagc cactttcctt gagaccaaga gagtgcggtt gggatggggg ggagagcagg 2040
gtcccgtctg acaatggcgg ctgccataat ccagtgtagc taatggctct g 2091
```

&lt;210&gt; 35

&lt;211&gt; 2888

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1802448CB1

&lt;400&gt; 35

```
ccggaggagg tgagcgccag gtcgccttcg cggccccggg acacaggcag ggacgcggga 60
gctgatgcgg ctggaccggc cggggaaaca gtattttctg gaagggggcc cctctgaagc 120
gggtccaggat cctgcacatg gcgctgaccg gggcctcaga cccctctgca gaggcagagg 180
ccaacgggga gaagcccttt ctgctgcggg cattgcagat cgcgctgggt gtctccctct 240
actgggtcac ctccatctcc atggtgttcc ttaataagta cctgctggac agccccctcc 300
tgcggttgga caccctcac ttcgtcacct ctaccagtg cctggtgacc acgctgctgt 360
gcaaaggcct cagcgctctg gccgcctgct gccctgggtg cgtggacttc ccagcttgc 420
gcctggacct cagggtggcc cgcagcgctc tgccctgtc ggtggctctc atcggcatga 480
tcaccttcaa taacctctgc ctcaagtacg tcggtgtggc ctctacaat gtgggcccgt 540
cactcaccac cgtcttcaac gtgctgctct cctacctgct gctcaagcag accacctcct 600
tctatgccct gctcacctgc ggtatcatca tcgggggctt ctggcttggg gtggaccagg 660
agggggcaga aggcaccctg tcgtggctgg gcaccgtctt cggcgtgctg gctagcctct 720
gtgtctcgct caacgccatc tacaccacga aggtgctccc ggcggtggac ggcagcatct 780
ggcgctgac tttctacaac aacgtcaacg cctgcatect ctctctgccc ctgctcctgc 840
tgctcgggga gcttcaggcc ctgctgact ttgccagct gggcagtgcc cacttctggg 900
ggatgatgac gctgggcggc ctgtttggct ttgccatcgg ctacgtgaca ggactgcaga 960
tcaagttcac cagtccgctg acccacaatg tgtcgggcac ggccaaggcc tgtgccaga 1020
cagtgtggc cgtgctctac tacgaggaga ccaagagctt cctctggtgg acgagcaaca 1080
tgatggtgct gggcggtccc tccgcctaca cctgggtcag gggctgggag atgaagaaga 1140
ctccggagga gccagcccc aaagacagcg agaagagcgc catgggggtg tgagcaccac 1200
aggcaccctg gatggccccg ccccgggggc cgtacacagg caggggccagc acagtagtga 1260
aggcggtctc ctggacccca gaagcgtgct gtggtgtgga ctgggtgcta cttatagacc 1320
caatcagaat acggtgggtg agaaggaacc agtgtttaca agtaatatca gaaagtga 1380
ggaaccagtg tttaacaagta ataccagaaa gttgccaaac ccttctctat cctctcgtat 1440
ttctgagttt ttgtccttcc cgaggagga ccctagttag agttgaaccc ctctctctg 1500
cctccagggc ctgtctgcct ccacatcact ctgaggacag ggacaggcaa caaccttga 1560
gggacagcaa ctggacccca gaagcgtgct cactgtactc aggggagatg gccctaccac 1620
agccacctgg agaggggttg gaagccttcc gtctgcggtg ggcaggcagc ctaccctgg 1680
gcccacagac cggccttccct ttggaggaaa gtgtggcctg gactgaggga ggaaatgagc 1740
gagttccctc tgacaccagc agatccccag gggctgctgg gcagtggcct gggaatgggg 1800
tggattgtga gaaagtgtc accatctata caccctgtat gtccagcttt tgaacacaag 1860
ggaacctatg ttctcttaga ggttaagcag ggtcattaac atctccccc agtccctaag 1920
atcacattgt cctgcgtggc tcctctggcc ctgagtggca cctgtccctc tggctctcca 1980
```

PF-0681 PCT

```

gcacctggcc caggtaacag ccttctgaaa gcagagccaa ggagctgctt ctctcttctc 2040
ccagttctac ctccccagaa gccttctccc ccaggtgggg ctgatggagc aagggtccag 2100
actaggagcc ttccacccca gctgtgtctg gcgccctag atctctgcaa gggaggtgtt 2160
acagctgggt ctgagccgct tgccttgtga tggtaagaca ccaaccttta cattcttccc 2220
tgaggttgtg gctgacagag cctgcttggc cccactgtta gtccagttag ctcttatatc 2280
aaaatgccgt agcgcgggtg gcttacaaac aacagaaacg tattgctcac agtcctgggg 2340
ggctgggacg tccaagatca agaggccagc agattcggac tccgctgagg gctgtttccc 2400
gatccataga tgggtgcctc tcgctgtatc ctcaatggta gaagcacaaa caagcaagct 2460
ccttctgccc tcttttataa ggactccaac cctgttcatg agggctctgc ccccatgacc 2520
caatcagctc caaaggcccc acctcctaact actgtcacct tgggggtgag aattccaatg 2580
tgaatttgca gggggagtgg gggacacaca caaatttcgg ggccatacca cccttcacca 2640
caccctcctg cgtcaggggt ggcttgcaat ccctggccct tctggtgggc atttggtatg 2700
tcctttctct tggggtgatt tctgatgttt ttactctata tagtgaaaag ctaggagag 2760
cgggtcttct ccccccctcc tctccagtcc cctcacatc ccagatgggt tctaattgag 2820
ctgctggggc ctgatgccct gagttgtttg tgattcaata aagaatccat aagaaaaaaa 2880
aaaaaaaa

```

&lt;210&gt; 36

&lt;211&gt; 2323

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2057214CB1

&lt;400&gt; 36

```

ccctcagcgg cgcttctctg tagcgagcct agtggcgggt gtttgattg aaacgtgagc 60
gcgacccgac cttaaagagt ggggagcaaa gggaggacag agccctttaa aacgagcgg 120
gtggtgcctg cccctttaag ggcggggcgt ccggacgact gtatctgagc cccagactgc 180
cccaggtttc tgtcgcaggc tgcgaggaaa ggcccctagg ctgggtctgg gtgcttggcg 240
gcggcggtt cctccccgt cgtcctccc gggcccagag gcacctcggc ttcagtcag 300
ctgagcagag tatggaagca cctgactacg aagtgtatc cgtgcgagaa cagctattcc 360
acgagaggat ccgcgagtgt attatatcaa cacttctgtt tgcaacactg tacatcctct 420
gccacatctt cctgacccgc ttcaagaagc ctgctgagtt caccacagtg gatgatgaag 480
atgccaccgt caacaagatt gcgctcgagc tgtgcacctt taccctggca attgccctgg 540
gtgctgtcct gctcctgccc ttctccatca tcagcaatga ggtgctgctc tccctgcctc 600
ggaactacta catccagtgg ctcaacggct cctcatcca tggcctctgg aacctgtgtt 660
ttctcttctc caacctgtcc ctcatcttcc tcatgccctt tgcatatttc ttcactgagt 720
ctgagggtt tgctggctcc agaaagggtg tcctggggcg ggtctatgag acagtgggtg 780
tggtgatgct cctcactctg ctggtgctag gtatggtgtg ggtggcatca gccattgtgg 840
acaagaacaa ggccaacaga gagtcactct atgacttttg ggagtactat cccccctacc 900
tctactcatg catctccttc cttgggggtc tgctgctcct ggtgtgtact ccactgggtc 960
tcgcccgcgt gttctcctgc actgggaagc tgctagtcaa gcccggctg ctggaagacc 1020
tggaggagca gctgtactgc tcagcctttg aggaggcagc cctgacccgc aggatctgta 1080
atcctacttc ctgctggctg cctttagaca tggagctgct acacagacag gtccctggctc 1140
tgcagacaca gagggtcctg ctggagaaga ggcggaaggc ttcagcctgg caacggaacc 1200
tgggctaccc cctggctatg ctgtgcttgc tgggtgctgac gggcctgtct gtgctcattg 1260
tggccatcca catcctggag ctgctcatcg atgaggctgc catgccccga ggcattgcagg 1320
gtacctcctt aggccaggtc tccttctcca agctgggctc ctttgggtgc gtcattcagg 1380
ttgtactcat cttttaccta atgggtgtcct cagttgtggg cttctatagc tctccactct 1440
tccggagcct gcggcccaga tggcacgaca ctgccatgac gcagataatt ggggaactgtg 1500
tctgtctcct ggtcctaagc tcagcacttc ctgtcttctc tcgaacctg gggctcactc 1560
gctttgacct gctgggtgac tttggacgct tcaactggct gggcaatttc tacattgtgt 1620
tcctctacaa cgcagccttt gcaggcctca ccactctctg tctgggtgaag acctcactg 1680
cagctgtgctg ggcagagctg atccgggcct tggggtgga cagactgccg ctgccctgtc 1740
ccggtttccc ccaggcatct aggaagaccc agcaccagtg acctccagct ggggggtggga 1800
aggaaaaaac tggacactgc catctgctgc ctaggcctgg aggggaagccc aaggctactt 1860
ggacctcagg acctggaatc tgagaggggtg ggtggcagag gggagcagag ccattctgcac 1920
tattgcataa tctgagccag agtttgggac caggacctcc tgcttttcca tacttaactg 1980
tggcctcagc atggggtagg gctgggtgac tgggtctagc ccctgatccc aaatctgttt 2040

```

PF-0681 PCT

```
acacatcaat ctgcctcact gctgttcttg gccatcccca tagccatggt tacatgattt 2100
gatgtgcaat aggggtggggg aggggcaggg aaaggactgg gccagggcag gctcgggaga 2160
tagattgtct ccttgcctc tggcccagca gagcctaagc actgtgctat cctggagggg 2220
ctttggacca cctgaaagac caaggggata gggaggagga ggcttcagcc atcagcaata 2280
aagttgatec cagggtttgc tttgtttttt taaaaaaaaa aaa 2323
```

&lt;210&gt; 37

&lt;211&gt; 1721

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2448177CB1

&lt;400&gt; 37

```
cccagaagct gggccctggc ggcggatctg gacgtggcgg cccgcgccgg gctgagtcac 60
caggaggagg ctgtggccga ggacgccgag gcctggagtg ggtggtagcc ccgagctggg 120
accctcctcc ctccacaatc tcccagggtc tgcaggagacc gagggcctac actgcctcct 180
cccacgccgt gcttaggaac ccttgcctgc ctccagaacac cagcgcctc cctccgggtg 240
agccctgcct gggcgggggc cctcctcca cagccatggt ctggaagaaa ctgggctccc 300
gcaacttctc cagctgcccc agtggctcca tccagtggat atgggatgtg ttgggtgaat 360
gtgcccagga cggctgggac gagggcagcg tgggcctggg cttgatctcc attctctgct 420
ttgctgcatc taccttccc cagttcatca aagcctacaa gacgggcaac atggaccagg 480
cgctgtccct gtggttccct ctgggctgga ttggcgagga ctctgcaac ctcatcggt 540
ccttcttgct tgaccagctg cccctgcaga cctacacggc tgtgtattat gtcttggcag 600
acctggtgat gctgacgctg tacttttact acaagttcag gacgcgcccc tctctgttgt 660
ctgcccccat caactccgtg ctgttggtcc tcatggggat ggcgtgcgcc acaccgctgc 720
tgagtgtgct tgggcccgtg gctgccccta ggggaagcctt cggggggcgg gcgctcctgt 780
ccgtggagtc gggcagcaag cccttcaccc ggcaggaagt cattggcttc gtcacgggt 840
ccatctccag cgtgtgttac ctgctttccc ggctgcctca gatccgcacc aacttctcct 900
ggaagtccac ccaggggatc tctactctc tgttcgcgct ggtgatgctg gggaaacacgc 960
tgtatgggct gagcgtgctg ctcaaaaacc ccgaggaggg ccagagcgag ggcagctacc 1020
tgctgcacca cctgccttggt cttgtgggca gcctgggctg gctgctgctc gacaccatca 1080
tctccatcca gttcctgggt tacaggcgca gcaccgcgc ctggagctt gagcccctcc 1140
tcccagctg accagaacca ggctgagcgc agggaggacag gcaccaccgg atgccacacc 1200
aggcaggagg aggtgtggac agtgatgta cggcgccct gcacagcct gcgggtggcc 1260
tctggatcct ccgtggaccg aaccgtcccc ccaggaacac accttcaggt agaccccgaa 1320
gcctcaaggc cggggctgga gcgagaccc cagggcctct caggagacag tgaggctgcc 1380
cctctacca cctacctcat tctgcctact caccacagg gccacagcca cagcctgctg 1440
gactcaggac tgtcctgtca actccagaca actgataaac aggcgggta cagtggctcg 1500
cacctgtaat cctagcactt tgggaggccg aagcgggtgg accacttgac gtccgtagtt 1560
cgagaccagc ctggccaaca tgggtgaaacc ccatctctac taaaaataca aaaattagcc 1620
aggtgtgggt gcacacatct gtagtcccag ctacgggctg caggatttcg atatcagggt 1680
tattcgatac gtcgaacctc aagccggggc ccggtaccca a 1721
```

&lt;210&gt; 38

&lt;211&gt; 2282

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2741701CB1

&lt;400&gt; 38

```
cggctcgagg ccctttgtga gggctgtgag ctgcgcctga cgggtggcacc atgagcagct 60
cagggtggggc gcccggggcg tccgccagct ctgcgcgcgc cgcgaggaa gagggcatga 120
cgtgggtgta ccgctggctg tgtgcctgt ctggggtgct gggggcagtc tcttgcgcga 180
tctctggcct cttcaactgc atcaccatcc accctctgaa catcgcggcc ggcgtgtgga 240
tgatcatgaa tgccttcac ttgttgctgt gtgaggcgcc cttctgctgc cagttcatcg 300
```

PF-0681 PCT

```

agtttgcaaa cacagtggcg gagaagggtgg accggctgcg ctcttggcag aaggctgtct 360
tctactgcgg gatggcggtc gttcccatcg tcatcagcct gacctgacc acgtgtctgg 420
gcaacgccat cgcttttgct acgggggtgc tgtacggact ctctgtcttg ggcaaaaagg 480
gcgatgcgat ctccatgccc aggatccagc agcagaggca gcaggcggat gaggagaagc 540
tcgcgagagac cctggagggg gagctgtgaa gggctgggag cccctccctc cctgtccctc 600
ttctggctct gtgtgggtcc aagtggggcc tggactgtcc acgtgaggc acagcctgga 660
gaggggcctt tgcacgtgtc cctacacctg gactcctctg ctctttctc cagactggct 720
taagccagga gccactggct gctgggtgtg gggctctggg tgctggactt gaggcagagc 780
ctgcagcagc tgtgtggaca ctacccagcc ctactcctct gctgggtggg tctgcagatc 840
tcacaccaca gacagggtc cctgtgacct gctgtgacct gggagcagct tccctggag 900
atgctgggtc gggttgagg ggaggggcaa gtgggacct gccacctggg cactgagcag 960
agggacctcc cccagctctc ttagcaggtg gagccccagg gcctgggaca gcctgccgt 1020
gccagcaacc tcccactgct gcctaggggt cagcgccccac tgtcaccttg ccttctgaag 1080
aagcccacag ggctcctaag gtgcaccccg gtacctgga ctgcagcctt ggcagtgtact 1140
ggagcagctg gtgggggatg ctccctgctg gccctgggaa ccttgagacag gccacctcaa 1200
ggccccctcg ctgccccctc tccctgggccc tgctggggcc cctaggttct gcccatcacc 1260
ccccgccctt gctggccttg gtgctaagga agtggggaga gcaggctctc cctggcaccg 1320
aggggtgccc cctctccctt ggtgtggccc cgtcaacatc agccacagcc cagccccatt 1380
agtgggttag tgggtctgac ctccagccca ctccaggtgt cctgtgtggc tgcccaagcc 1440
ctgcccctcag ggagcttctg ccttttaaga actgggcaga ggccacagtc acctccccac 1500
acagagctgt ccccactgcc ctgggtgcca ggctgtccgg agccaggcct acccagggtg 1560
gatgcagaga gctgggtgccc aggatgtgca ccccatatt cctctgccc tgtggcctca 1620
gcccgcgtgg ctctctgacc gtgaggctgg ctctcagcca tcgggcaggt gcctgtgtcg 1680
gcctggctta gccaggttg ggcttgagc aagcgggcg gtgtggaaga tattccatct 1740
ggggccaacc ccaggctggg cctgcgtgta gcttctggag cgcaggtact ggtcttctg 1800
aagtgaactg tttcccagga acacctctcg ggcccatctg cgtctgaggc tgggagtggt 1860
atctgaggcc gggagtggca tctgaggcca ggagtggcag gctgggtggc tgggcgtggg 1920
gttttctggg cctgcccag tactgccctg gggacttggg gggctcctgg gtcagcagca 1980
tcccaccctt gggagtctgg ccagctgagc ccagggttg caggggcatt atagcctggt 2040
ggacatgtgc cttcagggtt cctccggggc caccttctc aggccagtgc tgggttcaaa 2100
gggctgtgtg tgtgtgtgtg tgtgtgtgtg tatgtgtgtg ggtgcacaca 2160
tctgtcccat gtatgcagtg agacctgtct acctcccaca aggagcaagg gctctgccc 2220
ccctctgtct attcctaccc aggtagtggg accccggggc cccttctgcc tggcttgcc 2282
gc 2282

```

&lt;210&gt; 39

&lt;211&gt; 967

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3487228CB1

&lt;400&gt; 39

```

ggacagctgg cgccggcagc agcgggaccg cgggagacgg cggcttgcca gctggctgtg 60
ctgcaggagc tgcggcagcc gccgtccttg cactaagga acctctgtg ccagcccttc 120
ctcacttttg gaaatggcgg gtgaaattac agagaccggg gaactctact ctccctacgt 180
tgggctgggtg tacatgttta acctcatcgt gggcacgggc gactcacca tgcccaaggc 240
attcgccact gccgggtggc ttgtcagcct cgtcctgctg gtgttcctgg gttcatgag 300
cttcatgacc accacctttg tgatagaggc catggctgca gccaacgcgc agctccactg 360
gaagaggatg gagaacctca aggaggagga agatgacgac tcctccacag cctcagacag 420
cgatgtttct atccgggaca actacgagcg ggcagagaag cggcccatcc tgtctgtgag 480
taagtcttgg ggttctcgcg ggcccgact tccctccggg ccacagggtt ccctttcttc 540
atggagaggg ccctggagag gctccccgca gattctggca tttcctgccc ctgggttctg 600
aggcaggccc tgtggtagac tcaaaagagc acagggttcc agtcctggg acagtctgg 660
cccctgctgg cttggccact actcagattt gcagcttggg aagttccttg acctgtgcac 720
ctcagtttcc ccccataccc ctccctgtta ctgagatgat taaatgcctg gttccttgcc 780
tgggtgcaccc tggaaagctc agtagatgtc gttctgtggc atctctctt cctcctgcct 840
tgtgccctca tgttcataca tgccctgcc ttgtgtgtc ctccctggc ggctgggtta 900
ggctctgacg cctgctctcc ctttcacccc agcaggagca tctctgaatt cccttaaaaa 960

```

PF-0681 PCT

aaaaaaa

967

&lt;210&gt; 40

&lt;211&gt; 2221

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3671426CB1

&lt;400&gt; 40

```
agctgtgcga cctcttctgc ggccggcctg ggcagggtgc ttcctcgaga ggcaggcagg 60
ggatcccgga cgtgatcgtg cagttggcct tatacaggat gctgtgttct ttgctccttt 120
gtgaatgtct gttgctggta gctggttatg ctcatgatga tgactggatt gacccacag 180
acatgcttaa ctatgatgct gcttcaggaa caatgagaaa atctcaggca aaatatggta 240
tttcagggga aaaggatgtc agtcctgact tgtcatgtgc tgatgaaata tcagaatgtt 300
atcacaaact tgattcttta acttataaga ttgatgagtg tgaaaagaaa aagagggaag 360
actatgaaag tcaaagcaat cctgttttta ggagatactt aaataagatt ttaattgaag 420
ctggaaagct tggacttcct gatgaaaaca aaggcgatat gcattatgat gctgagatta 480
tccttaaaag agaaaactttg ttagaaatac agaagtttct caatggagaa gactggaaac 540
cagggtgcctt ggatgatgca ctaagtgata ttttaattaa ttttaagttt catgattttg 600
aaacatggaa gtggcgattc gaagattcct ttggagtggg tccatataat gtgttaattg 660
tacttctttg tctgctctgc atcgtggttt tagtggctac tgagctgtgg acatatgtac 720
gttggtacac tcagttgaga cgtgttttaa tcatcagctt tctgttcagt ttgggatgga 780
attggatgta tttatataag ctagcttttg cacagcatca ggctgaagtc gccaaagatg 840
agccattaaa caatgtgtgt gccaaaaaga tggactggac tggaaagtac tgggaatgg 900
ttagaagttc atggacctat aaggatgacc catgccaaaa atactatgag ctcttactag 960
tcaaccctat ttggttggtc ccaccaacaa aggcacttgc agttacattc accacatttg 1020
taacggagcc attgaagcat attgaaaag gaactgggga atttattaaa gcactcatga 1080
agggaaattcc agcgtgctt catcttccag tgcgtataat tatggcatta gccatcctga 1140
gtttctgcta tgggtgctgga aaatcagttc atgtgctgag acatataggc ggtcctgaga 1200
gagaacctcc ccaggcactt tggccacggg atagaagacg gcaggaggaa attgattata 1260
gacctgatgg tggagcagggt gatgccgatt tccattatag gggccaaatg ggccccactg 1320
agcaaggccc ttatgccaaa acgtatgagg gtagaagaga gatlttgaga gagagagatg 1380
ttgacttgag atttcagact ggcaacaaga gccctgaagt gctccgggca tttgatgtac 1440
cagacgcaga ggacagagag catcccacgg tggtagccag tcataaatca cctgttttgg 1500
atacaaagcc caaggagaca ggtggaatcc tggggaaggc acaccgaaag aaagcagtac 1560
tgaaagcagc cagtcggcca agcctgtctc tggccaagac acatcaggga atacagaagg 1620
ttcacccgca gcggaaaagg cccagctcaa gtcggaagcc gcaggcagcc cagaccaagg 1680
cagcacatac agccccgcaa gaggtgtggc tggaccacgt ggacaggatc cggtcagcag 1740
cccctgtggc tagaggaaca ccagcacaaa cgacagcctc aagtcctcct cgagctttat 1800
atccatttgg ggatgaagtc tactttgaca gctagcaagg cgacatgcaa ctgttgttga 1860
atgatgacag caattcagga aagacttaaa tatgaaagca aattgaacac atcggtgtgt 1920
tgttatcaga aaagagatga gatgagataa gacttgttta ttgactagcc aatatgtcat 1980
taaaaattaag gtttatattg tgaagctttt ttgaaatttt actttttttt gagagtcttg 2040
ctctgtcacc caggctggag tgaggtggta caatctcggc tcaactgcaac ctccacctcc 2100
caggttcaag agattctcct gcctcagctc cccgagtagc taggattaca gatgtgcatc 2160
accatgcctg gctaattttt gtatttttta gtagagatgg ggtttcacca tgttggtcag 2220
g 2221
```

&lt;210&gt; 41

&lt;211&gt; 2001

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 4020882CB1

&lt;400&gt; 41

PF-0681 PCT

```

ctcactagtc ggcggccgct tgggcgcact tgccgggtca ccttgtcccg gaggagaaat 60
ggcttccctg aggcgaagtgt aacctacatt cccagcccac cagcctgacg cccagccagg 120
gagagagtag catggatggc atcattgaac agaagagcat gctggtgcac agtaaaatca 180
gtgatgctgg caagaggaat ggtttaatta acaccagaaa cttgatggcc gagagcagag 240
atggtctggt gtctgtttac ccagcgcccc agtaccagag ccaccgggtg gggggccagca 300
cagtgcgggc cagcctggac agcagcagga gtgagccgat gcagcagctg ctggacccca 360
acaccctgca gcagtcagtg gagtcccgt accggcccaa catcatcctc tattcagagg 420
gcgtgctgcg ctectggggg gacggtgtgg ccgcccactg ctgcgagacc accttcacg 480
aggaccggtc gccaccaaaa gacagcctcg agtaccgga tgggaagtgc attgacctct 540
cagctgatga cataaaaatc cacacctgt cctacgatgt ggaggaggag gaggagtcc 600
aggagctgga gagcgactac tcaagcgaca cagagagtga ggacaatttc ctcatgatgc 660
ccccgcggga ccacctgggc ctcatgtct tctccatgct ctgctgcttc tggcctctgg 720
gcatacgagc cttctacttg tcccatgaga ccaacaaagc cgtggccaag ggggacttgc 780
accaggccag caccagctcc cggcgggccc tattcctggc agtgctgtcc atcaccattg 840
ggactggcgt ctatgtgggc gtggcgtggc cgtggccctc atcgccctacc tctccaagaa 900
caaccacctg tgagcttctt gcgaatggag ggggagcacc cggggccagg tctgtgtgga 960
cgtggaggaa gcaggcatac cgcatagtgc tgtacgtac aaatgattgc caaatgatgc 1020
cacgaagccc tgggatttcc taccatgga tttattttgt tttatcctt taatttcattg 1080
ttcacagcac tgtgtagagc accagacaga cgggactgc taatccttc aaaggaaagc 1140
tccaaagatc ccagccgca aggtgtctc tggatggatt ctggtggatg aatggcaacg 1200
cggctctctg cagcctgcca gtgcccagag tgccaccgca ttagcaatat acaaacagtc 1260
caaaaaagtg tttatttttt atggaatacg gtgcaatagg cagaggacaa gggacacatc 1320
actcttctgt ctgtggccct gctggagtc tttgtgcccc ccggagtcca cagccttcc 1380
ctgcaagacg agaatggggc tgggaagaaa gaggcaacac caggctggc agggagcccc 1440
ctgcactgct ctgcagaccc attggcctga ccctgagaag cagagccagc aaagccggg 1500
acctgcccct ctttctttcc cttcacacca cccagcctc aggatgtcaa gccacctccg 1560
gaacgtgtct aactccaca gctacccgc agcaatacgc actcttggga cctcgtgat 1620
ctaggatggg gaggcaggcc accgcccctc ccaagactcc tcaagaaaga gccccgcggt 1680
tgctccggaa actcgaggca ctgcagctat gggcactgcc tcagcctaaa gacacagggg 1740
cgctcccaa tcaccgcgt ggcggtgct caccctgca taagcagaaa ctagtgatcc 1800
tggaatgag atgggcctta ctctgtcgac taaatgaata gctattttct tgtcattttt 1860
taaagtgcaa ctctgtctc atgctgctta agttaccaga tgaatgctga gaaataagta 1920
atcacagaca ttttaatacc atttcattgc tgttttacga gtgttcatta cttacaaaaa 1980
aattatcttt taaaaaaaaa a 2001

```

&lt;210&gt; 42

&lt;211&gt; 1170

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 4069777CB1

&lt;400&gt; 42

```

gtcagcagtc ggggtctgtg tccgagtggg gatggctctg ggggctgcgg atactgtgca 60
ggccaagcgc tcggccgtct cgggcgtgct cagcatagtt ggcatttccc tggctctggg 120
cagcctgata agcatccttg aacatcagct ggggcatatt tttaccaatg atgagtagag 180
atgtcattgc cctggtgagc caggtcttgc cggtttatag tgtctttcac gtgtttgagg 240
ccatctgttg tgtctatggc ggagtcttga gaggaactgg gaagcaggcc tttggtgccg 300
ctgtgaatgc catcacatat tacatcatcg gctaccact gggcatcctt ctgacctttg 360
tggtcagaat gagaatcatg ggctcttggc tgggcatgct ggctgtgtc ttcctggcaa 420
ctgctgcctt tgttgcttat actgcccggc tggactggaa gcttgctgca gaggaggcta 480
agaaacattc aggcggcag cagcagcaga gagcagagag cactgcaacc agacctgggc 540
ctgagaaagc agtcttatct tcagtggcta caggcagttc ccttggcatt accttgacaa 600
cgtattcaag gtctgagtgc cactgggact tcttcaggac tccagaggag gccacagccc 660
tttcagctcc taccagcaga ctatcagtga aacagctggc catccgccgt ggggctgtct 720
tgggggcggc gtcagccaca ctgatggagg ggctcacggc caggatccta gccaccaggc 780
actagcaaaag aagcttgga atagaaagcc aggagtggct gtccccagta tgcaaacaca 840
ccacggctct ccttgcaaaa acaccaatgg ggtctagtgc aggtggacac tttgaaccac 900
tcctcaaaaa aaagaacttt ggctgattcc ttgtggtgac actcagaggg gtctgaacag 960

```



PF-0681 PCT

```
acttgacaat tctgttctgg tcaagctgga gttttcttct gtgacttga ctgctctaca 1020
gaagacatca gccaactgca cgagtcagag tccagggatt gtcactatta ttaataatgt 1080
aaatggcttc aaatgggaca ctgcagataa aatcacaaaa accactgtta tattaaagat 1140
tacacatttc ctgggcaaaa aaaaaaaaaa 1170
```

<210> 43  
<211> 1912  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 5309830CB1

```
<400> 43
gaccccggtt gacggggccc ggcggcgtta actggagcga accccagcgt ccgcccacat 60
ggcctggacc aagtaccagc tggtccctggc cgggctcatg cttgttaccg gctccatcaa 120
cacgctctcg gcaaaatggg cggacaattt catggccgag ggctgtggag ggagcaagga 180
gcacagcttc cagcatccct tctccaggc agtgggcatg ttctgggag aattctcctg 240
cctggctgcc ttctacctcc tccgatgcag agctgcaggg caatcagact ccagcgtaga 300
ccccagcag ccttcaacc ctcttctttt cctgccccca gcgctctgtg acatgacagg 360
gaccagcctc atgtatgtgg ctctgaacat gaccagtgcc tccagcttcc agatgctgcg 420
gggtgcagtg atcatattca ctggcctgtt ctgggtggcc ttctggggcc ggaggctggg 480
gctgagccag tggctgggca tctagccac catcgccggg ctgggtggtcg tgggcctggc 540
tgacctctcg agcaagcacg acagtcagca caagctcagc gaagtgatca caggggacct 600
gttgatcatc atggcccaga tcatcgttgc catccagatg gtgctagagg agaagttcgt 660
ctacaaacac aatgtgcacc cactgcgggc agttggcact gagggcctct ttggctttgt 720
gatcctctcc ctgctgctgg tgcccattga ctacatcccc gccggctcct tcagcggaaa 780
cctcgtggg acactggagg atgcattgga cgccttctgc cagggtggcc agcagccgct 840
cattgccgtg gcactgctgg gcaacatcag cagcattgcc ttcttcaact tcgcaggcat 900
cagcgtcacc aaggaactga gcgccaccac ccgcatggtg ttggacagct tgcgcaccgt 960
tgtcatctgg gcactgagcc tggcactggg ctgggaggcc ttccatgcac tgcagatcct 1020
tggttccctc atactcctta taggcactgc cctctacaat gggctacacc gtccgctgct 1080
gggcgcctcg tccagggggc ggcccctggc agaggagagc gagcaggaga gactgctggg 1140
tggcacccgc actcccatca atgatgccag ctgaggttcc ctggaggctt ctactgccac 1200
ccgggtgctc cttctccctg agactgaggc cacacaggct ggtgggcccc gaatgcccta 1260
tccccaaagg ctcaccctgt cccctccctg cagaaccccc agggcagctg ctgccacaga 1320
agataacaac acccaagtcc tctttttctc actaccacct gcagggtggt gttaccacagc 1380
ccccacaagc ctgagtgcag tggcagacct cagctctctg gaccctcct acagcataag 1440
agctaaatca tgaagttaa ttgttagaat ttaccaccgt agtgtatctg aatcataaac 1500
tagattatca tagttatcta gtttatgagt cataagctag atttgattct tcaaagaaga 1560
aaagtacgtg agcaaattha aaccagaact ccagcctggg caacagagca agactccatc 1620
tcaaaaaaga aaaaaaaaaa ttaaagggtt tggactggat gcagtggctt atgctttaa 1680
tcccagcact ttgggaggct gaggcaggcg gattgcttga gcccaggagt ttgagaccag 1740
cctgggcaac atagttagac tccgtctcta aaaaaattta taaaaatgac cgggtgtggc 1800
ggctcacacc tgtaatccca gcactatggg aggccaaagg aggtggatca cctgaggtaa 1860
ggagtctctc cctcctcagc ctcccaaatc actgggacta cctctgtgag cc 1912
```

<210> 44  
<211> 2022  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 306263CB1

```
<400> 44
agattactaa agaacaaaga ctacgaagag aaagccagta ccaagagaac cgccggcagtt 60
tggtagtaac aaacggcgga gtggaacatg atctggatca gattgactac atagacagct 120
gcaccgcaga ggaagaggag gccgaggtga gacagcccaa gggaccagac ccagacagcc 180
```

PF-0681 PCT

```
ttagttcaca gtttatggcg tatattgaac agcggcgaat ctctcatgag gggttcaccag 240
taaagccagt agccattagg gagtttcaaa aaacagaaga tatgagaaga tatttacatc 300
aaaacagggt accagctgag ccattctccc tcctgtcact atcagcaagt cacaatcagc 360
tgtcacacac agacctggaa ctctcatcaga gaaggagca gttagtagag cgactcggg 420
gagaggctca gcttgctgcp ctgcagtatg aggaggagaa aataaggacc aagcagatcc 480
agagagatgc tgtctggac tttgtcaaac aaaaagcatc acaaagtcca caaaaacagc 540
accgctcct agatggcgta gatggtgagt gccccttccc atccagaagg tctcagcaca 600
ctgatgatag tgccttgctg atgagtgatg acagacctaa tgctctatta agttcacctg 660
caacagaaac agttcatcat tcccctgcat attcttttcc tgctgctatc cagagaaatc 720
agcctcagcg ccctgaaagc ttcttttccc gagcaggtgt cagggcagaa accaacaag 780
gtcatgcttc accccttccc ccattctgctg cacctaccac tgattctaca gattccataa 840
caggacagaa ttcaagacag agagaagaag agctggaatt aatagaccaa ctgcgtaaac 900
atattgagta ccggttgaaa gtgtctctac ctgtgtatct cggagcagct ctaactgacg 960
gtgtgtttct ttgccatttg gccaatcatg tgcgacctcg atctgtccca agcattcatg 1020
ttccctcacc agctgtacct aaattaacaa tggcgaaatg cagggcgaat gtggaaaatt 1080
tcctagaagc ttgcagaaaa attggtgtac ctccaggaaa tctttgttcc cttcccgaca 1140
tccttcagct aaacctcagc gttaaaagaa ctggttgaaac gctcctttct cttggggcac 1200
actcagaaga atccagtttt gtctgtctct ctctgcagct tctgggtttt gtggcatttt 1260
agtgtactgt gatgttaact ctctgtgtgc tttattactg gctcttcccc gctcgctgaa 1320
agattgcact ccggtgactt tctgcctcat tcctgtgctt ggtaaaggag gagcaattat 1380
gtttgctctt ccacatttta gaagagaaag gtttgagcca gggtgcagtg acggtccagg 1440
ctctcctgga acttgcccca cccaagcaac agcagacca gttatctgct gtttgaggat 1500
ccccaggacg gtgggactg gcctggccaa aacaaggaa aggacaccgt gattgctgct 1560
gccagctatc tgcttaaca aagctcttgt gtgttctcag aagggaccgt tccatgatt 1620
cctaacagga atattttgct tcatttctcc attttaaggg aggtattttc catttccatt 1680
gaattttttt acgaagacac cattgggttt ctccagatgaa acttctctta ctgaaaaccc 1740
agcacctaac ttggctgtgt tttctggag agcaggcgtg tttcacgtgc cagaaagcac 1800
atcttctgtg ttaggcacia gcatgctgtg ccattgtttt tacctggtct caagacaact 1860
ggtctgctgg tttctaaact tggtgtttca ttagactctc ctggagtatt tggttaaaata 1920
cagattcaca ttcagtagag ctggggtaag gcttaggaat ctgcattttc aacaaccact 1980
gcactccagc actggcaaca gacaggatcg gagggcgga ac 2022
```

&lt;210&gt; 45

&lt;211&gt; 1877

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 483751CB1

&lt;400&gt; 45

```
tgacggccgc ctcccggcta ccggggactt ctggagtccg agaagtcaac ggcgcggttg 60
ctgcggccgc cgcgtcccc ggcccaggcg gatggagaac ggagcgggtg acagcccac 120
tacggaggag gacccgggccc ccgcccaggc ccccggagc ggcctcgctg cctacttttt 180
catgggcccgc ctcccattgc tccggcgcgt tctcaagggc ttgcagctgt tgctgtctct 240
gctggccttc atctgtgaag aagttgtatc acaatgtact ttatgtggag gactttattt 300
ttttgagttt gtaagctgca gtgcctttct tctgagtctc cttatactga ttgtgtattg 360
cactccattt tatgagagag ttgataccac aaaagtaaaa tcatcggatt tttatattac 420
tttgggaaca ggatgtgtgt tttgttggc atccatcatt tttgtttcca cacatgacag 480
gacttcagct gagattgctg caattgtgtt tggatttata gcaagtttta tgttcctact 540
tgactttatc actatgctgt atgaaaaacg acaggagtcc cagctgagaa aacctgaaaa 600
taccactagg gctgaagccc tcaactgagcc acttaatgcc taaagactct ggggagcaga 660
tgttacctaa ggtagtgaac ctgcattgtg gtgctgagc cctggcagaa gctcttgtaa 720
aatttgtaaa ttgtttaaac cacttctttt ggagagcaag gggaaggta agaaggcagt 780
tttatcaata ttgtgtcagt caccacaaag taggcagat aagttaaaaa aaattttttt 840
ttaaataata attgaaactt atctcaaatg gagatttttg tgggaggagg agaaaacaat 900
tgtttttaaa tcacacagct caacggttga taaatgatc tgctattctg ttacaggtca 960
ttcttttact aggccttagct tccaaattat gctttatagc tgtataaaca tcgtgattat 1020
atccatctac ttgaaattg ttttattttt aaattaattt gcttagctgt ttgttttgat 1080
gcttagatta tgttctgtta atgggaattt aacatattta agaaaccaat atttaaaatg 1140
```

PF-0681 PCT

```
ttgggtctagg ttttttttccct taacatatat taccaggcctt tactgtatatt cactcagcct 1200
taaagtgttat aatatttttttg gataacgggtt attaatctctt tgagaccttc gtatagccta 1260
taaaatgtat gggagatggtt ggtatttttat gtgtataaaa gcaacaatat cagcaacttc 1320
gtgttttatac tgcaccttgg ttgttgatgt caagtataaaa aaagattggt ttgtaacaca 1380
taaaaaaatg gaagaaactg ataccacacc taaggaccaa agataagaaa gacttttttgc 1440
ccaagacagt gaaagtaatt ataaaaacaa gctttgacca cttaccaagt atctgaagag 1500
atgagttcat actatgattt agaaagtgggt tcaattcccc tgttggcata tgattatttt 1560
tactaaaatt aatacagctc tgtgggtctt ccttagtggt ttctttgaag ccaatctgtt 1620
tttttttagga caccagcctt tggtttttca tctgttcgag atgcctcttc tctgtctcct 1680
tatcagatag aaatggagtc atgtgctgct gcttcatcta gcagagggtg gcctctggct 1740
ctgacacttt ttgtcagttg tcttttaggtg gtcttgaatc ttgggcccctt ttgattgtga 1800
atactgtgta gcaggatctt gagagtcctt gttcttacat aggcattgct ctagtttgtc 1860
tttggcaaaa aaaaaaa 1877
```

&lt;210&gt; 46

&lt;211&gt; 2382

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 727332CB1

&lt;400&gt; 46

```
tatcagcagt aagtgttttc atatttttgt tcatcaaaag tgaaatacca aataacacgt 60
gttccatgga taccactcca ttaataacag agcatgtgat caacacaacc caagaccct 120
gttcctgggt ggataaaactt tctacagtac accaccgcat agtgggctgc agtcttgag 180
tgatatctgg agtactctat ggatctacat ttgtgccaat catctacatc aaggaccaca 240
gcaaaagaaa tgatagtata tatgcagggg caagccaata tgatttagac tatgtgtttg 300
cgactttcag tggcatcttt cttacaagta ctgtctactt tctggcctac tgcatagcca 360
tgaaaaatag tcctaaacta tatcctgaag cagtcctacc aggattcctg tcaggagtac 420
tttgggctat agctacctgc tgttgggttca tagcaaatca ctctctgagt gctgtgggtca 480
gttttccaat aatcactgct ggtccaggat ttatagctgc aatgtgggggt atcttcatgt 540
ttaaggaaat aaagggtcta caaaactacc tattaatgat acttgcattt tgcattcatc 600
tgactggagc cttatgcact gctttttcta aaatctaaca atgacaaaac cagcagggtg 660
cagcagtagt taagagaacg cgtctatcgg acagcggaga gatcatgctg agaaaagagt 720
gcattttcat atagcaaatg gatctcagcc actgttggag tgggtaaatg atttttttcc 780
ccaaaaatgt gagaatgaag gaagattgag ttctattcaa gtataaaaaat gaaaatttct 840
tctgatgaac tgtttatgaa ggtagtactt tactgggtga ctaaaatgtc tacattatta 900
tagcattaca tacgaggatg ctctttttac cacagtatgt acctagtgtt gcataatcta 960
gaacacagta tagagtccat atacaaagaa gttacagtcc atatttctca cctattaccc 1020
cagtgaagtg aggcagaaat accttgggag agggaggaga gccttacctt ctaatggaat 1080
tagattccaa agatatttgg atccttttgt tcataaagta aatgcatagt tggtaactaa 1140
atgttaaaaa ggccatttta ttattgttag ccaaaacttc tgtttatttc aagttaatat 1200
cgcccaactc tacatcatct acccatttta gtcataatg cttttgatta tcgccagata 1260
aagtccattc tcagagaaca aaaatttttc atcattgaga aaatgtgcat aacataatac 1320
atacatacac aagtatatcc acgtacataa atacacacag atatgttgct aaagtgagat 1380
tgtaggttat cagtaacaat tttaaaggaaa caaccactct tttgaatgta caactttttt 1440
ttctgaaata cactttttcta aaacctctt aactggagggt ttctctgttt cttcttgttt 1500
taactgggtg ttccagtag gcaaaaataaa atgacaggca tgctagtttt tttttttttt 1560
ttttttttt aagaaaactg aaattaagcc aggcagggtg gcgcactact ataattcccag 1620
ctgctggggc ggctgaagta ggagggtcac ttagcccaag agtgcaaatc tgccagccta 1680
ggcaacatag caaggtccca tctctaaaaa agaaatttaa tttaaatttg taagaaaaaa 1740
gaaaactgaa acagaactgc cagtacaatg tttgttgcaa ttgtttccaa aactttgaaa 1800
taacgaaact cctattcaaa tgtaaatttt acccattccc acctgatagt actgtttatt 1860
actttgtcat gcaggagaaa caaaacagct ttttatttgt gtgtatactc tgaagtcaat 1920
gtgcatgggt taaatcccac ctctgcttca ttttatttgt gtggctttgg acaagtctct 1980
aagccctctg tgtcaggccc ctcatataca aaatggagtt tatgatggca tctacctcac 2040
atggttattg tgtgggtgaa attagttcaa atctgtaaag tatgttgatg agtatctgggt 2100
acttggcaaa cactatataa ttgttagctt tattattatg tgttgtttga agggaaaaaa 2160
agtttttcta gattggagaa caccatgtaa gagctactaa gctcaccacc ttcccatcca 2220
```

PF-0681 PCT

```

aacaattatt actcctctaa atgaatacta cttaaaagtg atggagcttc tgtttgttta 2280
ttgaattgct aatacaataa aatatttgac tataacttct tataactcatt ttaatttttaa 2340
cctgaaatgt cattaaatta tgctgataaa ctttgaataa at 2382

```

&lt;210&gt; 47

&lt;211&gt; 2383

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 778800CB1

&lt;400&gt; 47

```

cttacaatca tcctgcttct tgtacatcac tgaattttca tctcattgat tatcatctgg 60
cagcattcat cacagtgatg cttgcgagga ggcttgatg ggctctcatc tcagaggcta 120
ctaaggcagg tgcagcatca atgattcact acatggttct gatatcagct cgcttggtac 180
tactcacttt gtgtggatgg gtactttgtt ggaccctcgt caatctcttt cgaagccatt 240
cagtcctcaa tctccttttc cttggctacc cgtttgggtg ttatgttcct ctttgcgtgt 300
ttcatcaaga tagtagagca catcttcttc tcacagacta caactatgtg gttcagcacg 360
aggcagtaga ggaaagtgcc tgcactgtgg gaggttggc caaatccaaa gactttctct 420
ccttgttgct ggagtcgcta aaagaacagt ttaataatgc cacacccatc cccaccaca 480
gttgtccctc atctccagac ctcatctgca atgaagtaga atgtctgaaa gcagatttca 540
accacagaat caaggaggtt ctcttcaact cctcttcag tgcctactat gttgcatttc 600
tccccctgtg ttttgtgaag agtaccagct actatgacat gcgctggtca tgtgagcacc 660
tcattatggt gtggatcaat gcttttgtea tgctcaccac gcaactgttg ccatccaaat 720
actgtgattt gctacataaa tcagctgctc acctgggcaa gtggcagaag ttggaacatg 780
ggtcctacag caatgctcca cagcacattt ggtcagaaaa tacaatatgg cctcaagggg 840
tgctggtgcg gcacagcaga tgtttatata gagccatggg gccttacaac gtggcagtcg 900
cttcagatgt atctcatgcc cgcttttatt tcttatttca tcgaccatta aggctgttaa 960
atctgctcat ccttattgag ggcagtgtcg tcttctatca gctctattcc ttgctgcggt 1020
cggagaagtg gaaccacaca ctttccatgg ctctcatcct cttctgcaac tactatgttt 1080
tattttaaact tctccgggac agaatagtat taggcagggc atactcctac ccaactcaaca 1140
gttatgaact caagacaaac taagctgcct ctcaacaatg agggagaact cagataaaaa 1200
tattttcata cgttctattt ttttcttggt atttttataa atatttaaga tattttatat 1260
tttgtatact attatgtttt gaaagtcggg aagagtaagg gatattaaat gtatccgtaa 1320
acaaggatgat gtgatctttc atgtgttagt atattcgaat aatatacata tgagaggtat 1380
gcctgtccag taaagagatt gatgtgtttg tatggctctg gaataactca ttcgtttatg 1440
aacacatctc atgtattact tgacacagta cctttgggtat gtctgctttt ctttcaatcc 1500
tcattatcct tagccctcag ctggccagta cttattgcta aagaaaggga gatcaatgga 1560
gttgccctacc agcaccatt ctcttttagt tcagctggct aatagttatg gaatacattg 1620
ggaaaagagt ggggttaaagt cctaggtggg gttctctttg ttgcctacca ggagaaagca 1680
gtaggcttcc tttctaagat acggactcag ctggctctatt ttaccctgtg acattttaat 1740
atacttttga ataagctact atcctgagge tgagggtttt gctgccactt ttctgtgtt 1800
tccacttctg cttccaaaga ttttctctt agttgtttga aataatactt tgccagacag 1860
gaagtcttta gactgccac attcatgaag aagatgatgt gtggctggat cagtaggaga 1920
agctgagtgc aaatgccatg tgggctagag tatgggaagg ttgaaaattc caagtaactg 1980
ggagagtggg caagtcctca gctgaacaat cttggatgta gtctagcagc aaccaccta 2040
caatggcggt atgtttggag ttatgagggg aagtgaaga taagaattac ctacagcttt 2100
gtactcagat aatttatatc caaaatttgt gttttaaatt tcatactcca gaactttcat 2160
gtttatgagt attagccttt taagttatgg gaggtgatat acttagtcca ttgatgttgt 2220
cacttgctcat ctaaacttat gagaccttgt cttagagaat ctaaaatata ttgtcatatg 2280
ttttctgtgc tgagtttgggt ttgttacact gtgaagcttt tgggtgaaaca aaataagact 2340
tatgaagtaa aaagaccaat cgtgtcgttg aaaaaaaaaa aaa 2383

```

&lt;210&gt; 48

&lt;211&gt; 1344

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

PF-0681 PCT

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1396995CB1

&lt;400&gt; 48

```
cacgcgtccg cccacgcgtc cggcagcatg tatgtgtcca gtcttcatta tcaactatat 60
gacccaaaga aagtaccaca tagtggtagt aaatattgat ttcagccacc aacagactca 120
tgtattcttc tagatgtcaa aggaaaattt tgtgaggacc actgtatatt acagcactct 180
aaccggtaag caaattggga attttaaatt atcaaaaatt ttttttttcc tgggtattgta 240
gcctgggatg gtcttgtggg cttggattga acaaagtgtc actaaacaaa catgcattaa 300
ttatcttcta ctgtgttgaa atggtattat tataggagtt tgtttggcaa gaatgagtat 360
tccagtgcct ggccactctg aacacataaa tttgattact gcaacatcca aatggaggac 420
cacatcatca gctgctactt gaagtggcca gtctgtgccc aactactgaa ctgttcagag 480
cccacagaaa gggctacggg agagacatgt atggtatctc gtatacacag gcatgcctta 540
tttctttgtc ttctacccca gtgctaccta cctaaagacc tctggtactc ttccctttcc 600
tgggaaagaa attcaatctt gttttctctt ttattattgt attctctatt ctacagttaa 660
gccccaatag tagccaaagg aaaaaggggg aagatgatag tttgactaaa caaataccag 720
ttcatacatt cttgttccaa taggagttat gggaggaaaa attattccaa gcaatcacac 780
aggatcaact ttttctcctt gtagaatatg tgtcatcaca ttggcagaat ctcatccagt 840
tcttcaaagt ggaaaaacaa ttaaaagcat ttctatcag atcagtacca actgtagcag 900
acttcagaac aaggctctctg ggaaatttaa gcgggtatgt tcctgtaatt ctgtgattat 960
ctacttacat cctctatatt catggtaatg acagatccaa tgaactttag aatccagtag 1020
catatgctta gcatacttct ctagcagttt gaggtgctaa ttttaggtat actttcacct 1080
aaagaaattc tcagctcccc caaattaggt atctcaggag gtgtagtatc tgttatatta 1140
ggttctgtgc tactatccct ataatgcccga ggatggagga gggggaaggc aggcctttga 1200
aaggagaact ctaatagcaa tataacaaga tattttgtcc cttctagttg tattaacta 1260
gtatgtcgag ttctgtaaaa tttaggtgac taactcttct ttaccatata tttccacatg 1320
taaattaaac agaaaaaaaaa ttgc 1344
```

&lt;210&gt; 49

&lt;211&gt; 800

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1597730CB1

&lt;400&gt; 49

```
gcgctggaag ccctggcggc ggccggcccat gggggccttg gcgctgcccg cctggctgca 60
gccaggtat aggaagaatg cgtatctttt catctattac ttaatccagt tctgtggcca 120
ctcttgataa ttacaaata tgacagtcag attcttttca tttggaaaag attcaatggg 180
tgacactttt tatgtatttg gacttgtgat gcgactttgc caatccgtat ctctcctgga 240
actgctgcac atatatgttg gcattgagtc aaaccatctt ctcccaagggt ttttgcagct 300
cacagaaaga ataatcatcc tttttgtggg gatcaccagt caagaggaag tccaagagaa 360
atatgtgggtg tgtgttttat tcgtcttttg gaatctattg gatatgggta ggtacactta 420
tagcatgtta tcagtcatag gaatcctcta tgctgtcttg acatgggtca gtcaaacact 480
atggatgcca atttatcctt tgtgtgttct tgctgaagca tttgccatct atcaatcgct 540
gccttatttt gaatcatttg gcacttatcc caccaagctg ccctttgact tatccatcta 600
tttcccatat gtgctgaaaa tatatctcat gatgctcttt ataggtatgt attttaccta 660
cagtcactta tactcagaaa gaagagacat cctcggaatc tttccatta aaaaaaagaa 720
gatgtgaagt acagcattcc agtgtgacac gagaaaagac aggctgtgga ttcagtgcag 780
taaataaaac acgaggaagg 800
```

&lt;210&gt; 50

&lt;211&gt; 1505

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1629304CB1

PF-0681 PCT

&lt;400&gt; 50

```
ggcggcgggg cgctccgggg agggcaggac agctgatggg tgtggcagaa acatctcaag 60
gtagctgggc cgccccact tccccatcta cctcttgtcc tcccccaac accaccacca 120
ccctggctcc cctccctcat gaccgcctgg atcctcctgc ctgtcagcct gtcagcggtc 180
tccatcactg gcatatggac tgtgtatgcc atggctgtga tgaaccacca tgtatgccct 240
gtggagaact ggtcctacaa cgagtcctgc cctcctgacc ctgctgagca aggggggtccc 300
aagacctgct gcacctgga cgatgtcccc ctcatcagca agtgtggctc ctatccccc 360
gaaagctgcc tcttcagcct cattggcaac atgggtgctt tcatgggtgg cctgatctgc 420
ctcctgcgct acgggcagct cctggagcag agtcggcact cttgggttaa caccacggca 480
ctcatcacag gctgcaccaa cgctgcgggc ctcttgggtg ttggcaactt tcaggtggat 540
catgccaggc ctctgacta cgttggagct ggcgtggcct tccctgcggg gctgctcttt 600
gtttgcctgc actgtgctct ctccaccaa gggggccacc ccccgctgga cctggctgtg 660
gcctatctgc gaagtgtgct ggctgtcat gcctttatca cctggctcct cagtggagtc 720
ttctttgtcc atgagagttc tcagctgcaa catggggcag ccctgtgtga gtgggtgtgt 780
gtcatcgata tctctatctt ctatggcacc ttcagctacg agtttggggc agtctcctca 840
gacacactgg tggtgcact gcagcctacc cctggccggg cctgcaagtc ctccgggagc 900
agcagcacct ccacccacct caactgtgcc cccgagagca tcgctatgat ctaaggtctg 960
gggaggggtg ctggccgggc ctccacagca cccacccca tatcttcttt ccatttattt 1020
cgtaccacaaa acaattttga gaaagtattc tgttgggac tgggcttctt cacttctgga 1080
gaagtggcca tcccatgccc acctgtgcca tggaggagtg ggccctgcca gctgccacag 1140
ctgcatgacc tgcttcccca cccacggtg tcgttttgtt tttaaaggtc acctgtcctc 1200
actcaccag ccagcccttc aggtgccttc tactcccagt gccaaagcca gaccactggg 1260
gtttcctgct gcaggaattg ggggctggga acagcagagg ggatagaagt ctggtggag 1320
tggagtgggc acgccttagc ctacggaaag gccatttct gggccactg agctgcactg 1380
ggattcttca ctctgcccct cacttctttt agggcaaata acacagcaga accacgtggg 1440
tatttagta ctttttttta tattaaga attctaattt gcaaaaaaaa gaaaaaaa 1500
aaaaa 1505
```

&lt;210&gt; 51

&lt;211&gt; 2744

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1989585CB1

&lt;400&gt; 51

```
gcacagtgag cggacgcctg ggcggcgggc gggcgggcgt gatggctccg gccgcggacc 60
gagaggggcta ctggggcccc acgacctcca cgctggactg gtgcgaggag aactactccg 120
tgacctggta catcgccgag ttctggaata cagttagtaa cctgatcatg attatactc 180
caatgttcgg tgcaattcag agtgtagtag acggtctgga aaaacgggtac attgcttctt 240
atttagcact cacagtggta ggaatgggat cctgggtgctt ccacatgact ctgaaatatg 300
aaatgcagct attggatgaa ctccaatga tatacagctg ttgcatattt gtgtactgca 360
tgtttgaatg tttcaagatc aagaactcag taaactacca tctgcttttt accttagttc 420
tattcagttt aatagtaacc acagtttacc ttaaggtaaa agagccaata ttccatcagg 480
tcatgtatgg aatgttggtc tttacattag tacttcgac tatttatatt gttacatggg 540
tttatccatg gcttagagga ctgggttata catcattggg tatattttta ttgggatttt 600
tattttggaa tatagataac atattttgtg agtcactgag gaactttcga aagaaggtag 660
cacctatcat aggtattacc acacaatttc atgcatgggt gcataatttt actggccttg 720
gttctatctt tcacatcctt ttcagtttgt atacaagaac actttacctg agatataggc 780
caaaagtga gtttctcttt ggaatctggc cagtgatcct gtttgagcct ctcaggaagc 840
attgatgaat cattccacca agaaaacaaa caagcaccta ccatagacct ggcagaataa 900
ataaggaaat ccttaaagat ctacaagttc aaatatgtca tgaccatcac agcagaggag 960
tgactttctg actaatgctg cccaccacac agagaataag gagtagggcc tgctgggtgt 1020
ttagctcatg gctttatctt atttgtcccc ctctcctttt cagctccag tttataaaga 1080
aacagagatg atgtgtgtgt atgcctcaa tgcagaaaca gttgggcttt tcttaacagg 1140
ttaacagttt gtgctggtta taagaaatta accttctttt ctttttgcca agggttgcat 1200
gtgaattata cgtttcttaa tatttgatta aataatgcag tcagctaaga gccacaaaa 1260
ccccatatcc agatagtag ctgtgagcca ccaatcagct gtgtcagatt tgattaatat 1320
tgtattaatc tcatggatgg ttctcaaagg taaaaaaga gagagaaaga gagagagcta 1380
```

PF-0681 PCT

```
gctgtgggat tctgtgtagt tcttcactat ctgttccagg gctagtcgga ggatcataag 1440
caagctcaga cctgccccca gttcagtact tgaccacatc ccaggattta tgggtgtctgt 1500
ctatggccac agccaaccag aatgagtggg tcatcttata accttcttga agctccatga 1560
ggacatgtat agaatgctgc accttcagcc agctctgctc tgatctttgc agatctttaa 1620
gttgtcactg tttctacggc tagacaaata tgactgggtg tctagactac ccatagttca 1680
tcctcctcac ataacttatt tgaagttgtc agggaaatgtc attacactct gattttttaa 1740
attatgcagt attttccagt tctcagagag gtcattgggtg tgcccaggga agtgtgttgc 1800
attaactaaa gctaaaatta tcagaatcct atttacataa agcatattaa aatatctgct 1860
ccttaatttt ctgaaggaaa tagatgtaag tttctgctac ttgaaacctc cagaaccgac 1920
aaaaattagg gtgctaaaaa aatgacagtg aaatctcatc gtaaggcagc tagttctaata 1980
ttgcatttga atgtataaca aattgcatca cttttcacaa acttaacacc tgcctggagt 2040
aaaaatctga tttgacctct ctgctcatgt ttcactatat ttgtagctct gcatcattta 2100
gcatttttat gacttttttt gctttatata aattattgta aaatccagat tgtttttacc 2160
acatgatgac gtgactcagg attaaatcct gagctacagt ttcaaaagga aaatcatttt 2220
tttcttatgc aaagatggta aaacactgtt gtcatttttt ccagatttca ataagtagcc 2280
tttgtacaaa tttaaaacca aaaactattc tgttgttttt cttcacctac tagccttagt 2340
ttttaaccaa gtatccaagg atgtcaactt gacatccttc ccttagcatg ccaacaagcc 2400
attagactga ttcttccacc tacttttatg aacctccagg aagaaaaaca cttgccaaaa 2460
cattaagaga actttgattc tgccttaaga gggataaaaa ctaaaccatt tgaaggcaca 2520
gcccataaga ctatacttga ttttggcccc aaggattctg acttctggct acatttcacc 2580
tgcagttacc tgggctatcg ggcgtgatac attttttaat tcattttgtt ttccaccatt 2640
ataatttaat acttttcagt tagatatacc tctgcattaa gaggtttata cagtgtgaagg 2700
ttaaggcgta tcggacatgg ccattgccgtt aaacatggga agac 2744
```

&lt;210&gt; 52

&lt;211&gt; 1679

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2134729CB1

&lt;400&gt; 52

```
tgactttggg ggcaccatgg accaaagtgg gatggagatt cctgtgaccc tcatcattaa 60
agcaccgaat cagaaatata gtgaccagac tattagctgc ttcttgaact ggaccgtggg 120
gaaactaaaa acgcatctat ctaacgttta ccctagcaaa ccattgacga aggatcagag 180
attggtgtat tcgggcagac tgcttcccga tcatctgcag ctgaaagaca ttctcagaaa 240
acaagatgag tatcatatgg ttcatctagt atgtacttct cggactcctc ccagttctcc 300
aaaatccagc accaatagag aaagtcatga agcattggca tccagcagca attctagtct 360
agatcattca ggatcaacaa ctccatcatc tgggtcaagaa acctgtgtct tagctgtggg 420
ttcttcctca gaaggattga ggcagcgtac ccttccacaa gcacaaactg accaagcaca 480
gagtcaccag tttccatatt taatgcaagg aaatgtagac aaccaatttc ctgggcaagc 540
tgctccacct ggattcccag tgtatcccgc gtttagccca ctgcagatgc tatgggtggc 600
acagatgtat gctcatcagt attatatgca gtatcaagct gcagtttcag ctccaggccac 660
atcaaattgc aaccaaccc agcctactac ttcacagcct cttaaatttg cacatgttcc 720
tggaagaaga ccccccaccg ctccaaacct agtggcccaa gaaaatcgac ccatgaatga 780
gaatgttcaa atgaatgcac agggagggtc agtactaaat gaagaagact tcaatcgaga 840
ctggctagac tggatgtaca cgttctcacg agctgcgatt ctcttagca ttgtatactt 900
ctattcttct tttagtccgt ttatcatggg aatgggagcc atgctactgg tttatttaca 960
ccaagctgga tggtttcctt ttaggcaaga agggaggtcat cagcaggctc ccaacaataa 1020
tgccgaagtt aacaatgatg ggcaaaatgc aaacaacttg gaacttgaag aaatggagcg 1080
tcttatggat gatgggcttg aagatgagag tggagaagat ggaggtgaag atgccagtgc 1140
aattcaaagg cctggattaa tggcttcagc ttggtcttct atcaccacct tctttacttc 1200
actaatacca gaggggcctc ccagggttgc caattgacct gaaaaactgt gccagctaca 1260
aggaggggtc gacttcagga aagtgggtta aataacagtg caatttcaaa aaaatttata 1320
actttctttt gatcatcatg tacagaggtg ttttttttct ttaggcttct catgcatatg 1380
aatattttta gcacgaatgg actactaaat atctgagttt tttttttttt tttttaaaga 1440
tcctaacaga acatagcgta acaatatttg tcttccagggt gttactcatt tcaattattg 1500
gtagtatacc aggacagacc tattttcatg tcttatttct ttaaagagct gcttcattgg 1560
ccgggcgcca tggctcacgt ctgtaatccc agcactttgg taggctgagg tgggtagatc 1620
```

PF-0681 PCT

acctgagggtc aggagtttga gaccagcctg gccaacgtga tgaaattcca tctctacta 1679

&lt;210&gt; 53

&lt;211&gt; 1625

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2299506CB1

&lt;400&gt; 53

```
gcgctcccca ccttttagcga gaccaacgag agaacaccgc ctgcagctag aacagcctgg 60
tcaggagcgt aacggagtggt tgcgccaacg tgagaggaaa cccgtgcgcg gctgcgcttt 120
cctgtcccca agccgttcta gacgcgggaa aaatgctttc tgaaagcagc tcctttttga 180
agggtgtgat gcttggaagc attttctgtg ctttgatcac tatgctagga cacattagga 240
ttgggtcatgg aaatagaatg caccaccatg agcatcatca cctacaagct cctaacaaaag 300
aagatatctt gaaaatttca gaggatgagc gcatggagct cagtaagagc tttcgagtat 360
actgtattat ccttgtaaaa cccaaagatg tgagtctttg ggctgcagta aaggagactt 420
ggaccaaaac ctgtgacaaa gcagagtctt tcagttctga aaatgttaaa gtgtttgagt 480
caattaatat ggacacaaat gacatgtggt taatgatgag aaaagcttac aaatacgctt 540
ttgataagta tagagaccaa tacaactggt tcttccttgc acgccccact acgtttgcta 600
tcattgaaaa cctaaagtat tttttgttaa aaaaggatcc atcacagcct ttctatctag 660
gccacactat aaaatctgga gaccttgaat atgtgggtat ggaaggagga attgtcttaa 720
gtgtagaatc aatgaaaaga cttaacagcc ttctcaatat ccagaaaaag tgtcctgaac 780
agggagggat gatttggaag atatctgaag ataaacagct agcagtttgc ctgaaatatg 840
ctggagtatt tgcagaaaat gcagaagatg ctgatggaaa agatgtattt aataccaaat 900
ctgttgggct ttctattaaa gaggcaatga cttatcacc caaccaggta gtagaaggct 960
gtgtttcaga tatggctggt acttttaatg gactgactcc aaatcagatg catgtgatga 1020
tgtatggggt ataccgcctt agggcatttg ggcataatct caatgatgca ttggttttct 1080
tacctccaaa tggttctgac aatgactgag aagtggtaga aaagcgtgaa tatgatcttt 1140
gtataggacg tgtgttgta ttattttag tagtaactac atatccaata cagctgtatg 1200
tttctttttc ttttctaatt tgggtggcact ggtataacca cacattaaag tcagtagtac 1260
atttttaaat gaggttggtt tttttcttta aaacacatga acattgtaaa tgtgttgga 1320
agaagtgttt taagaataat aattttgcaa ataaactatt aataaatatt atatgtgata 1380
aattctaaat tatgaacatt agaaatctgt ggggcacata tttttgctga ttggttaaaa 1440
aattttaaca ggtcttttagc gttctaagat atgcaaatga tatctctagt tgtgaatttg 1500
tgattaaagt aaaactttta gctgtgtggt ccctttactt ctaatactga tttatgttct 1560
aagctcccc aagttccaat ggatttgctt tctcaaatg tacaactaag caactaaaaga 1620
taattt
```

&lt;210&gt; 54

&lt;211&gt; 1212

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2506558CB1

&lt;400&gt; 54

```
actggtcaag ttcaagtcac acaggtttgc tgactgcgcc atattgttgc tgacacaact 60
ggagactgga cttaggaatg tttttgccac acttaacaga tgtccaaaaa gactcctgac 120
tgctgagtca acagctcttt ataccacctt tgatcaata ttggcaaaac acttgaatga 180
tggtaaaatc aatcagcttc ctcttttctt tggagagcct gctatggaat ttctctggga 240
tttctgaac catcaggagg gtccccgcac aagagatcat ttaagccacg gggagatcaa 300
cttacatgaa ttttcaaaag aaacaactaa tcagttgctt gcattttctc ttgtactgct 360
actcagattc gttgatgact gtctgtatc agtttttaag gaaaaatcag ccgtagaatt 420
gttgattagt cttgcagaag gctatagttc tcgctgtcat ccggtttttc agcttaaaaa 480
acaggtgctg agctgtgagg agagcatcag ggtttgggct ctgctgcctt tccccgaaga 540
actcactcgg caagccgtca gattagaaga taattctgaa acaaatgcct gccactcttt 600
```



PF-0681 PCT

```

gattacaaaa atgacggatg agctgtatca ccatatgcct gagaatcggt gtgtgttaaa 660
ggacttgat cgtcttccta ctgagacgtg gccccagctt ctccgtgagc tctgcagcac 720
acctgttccc accctgttct gccccaggat tgtgtgtgaa gtgctggttg tgctccgaag 780
cattggcgaa cagtgccgcc gtgtgtccag ccaggtcacc gttgcctcag agctgagaca 840
caggcagtgg gtggaaaagg cgctgcggtc tcgccagcgg cagaactacc tgcgtatgtg 900
gagtagtatc agactactgt cccctgtgct cagcctgata ctgttactca ttgctgtgga 960
gttggtaaac attcatgctg tttgtgggaa gaatgcgcat gagtatcagc agtacctaaa 1020
gtttgtaaaag tcgatcttgc agtacacgga gaacctgggtg gcttacacca gttacgaaaa 1080
gaacaagtgg aatgaaacta tcaatcttac acatacagct ttgttgaaaa tgtggacttt 1140
tagtgagaag aaacaaatgt taatacattt agccaagaaa tccacaagta aagtactctt 1200
atgaaaaaaa aa 1212

```

<210> 55  
 <211> 1465  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2546025CB1

```

<400> 55
cctttccagt ggacctgggc tgttgttgcg gttgttttcc ttctctccgt gcaacgctgg 60
caagtctcaa agtcgccaca gaaacatgcc cctgattcag tgccctctgct tagctgtaac 120
atgtaataca gaactacctg gcattcttcc gaacaagact ttcaataggg gccagtatgc 180
ttcgcttcat ccagaagttt tctcaagcat cttcaaagat actgaagtac tctttcccag 240
tgggactaag aaccagcaga acagatatac tttctctcaa gatgtctctc cagcaaaact 300
tttcccctatg tccaaggcct tggctttcct catcatttcc agcgtatatg agcaagacac 360
agtgttatca tacatcccc tgcagcttta aaaagcagca gaagcaagca cttctagcca 420
gacctcaag caccatcact tacctaactg acagcccaa gccagcatta tgtgtaactc 480
tggcaggact aatccccttc gttgtctcac cactggatcat gctgatgaca aaaacttata 540
ttcccatatt agcttttact cagatggctt atggagccag tttcctatct ttcttgggtg 600
ggatcagatg gggttttgct ctaccagaag gtagtccagc caaaccagac taccttaatt 660
tagctagcag tgcagtcctt cttttctttt catggtttgc cttccttatt tctgaaagac 720
ttagtgaagc catagtcaca gtaataatgg gtatgggagt agcattccac cttgaacttt 780
ttctcttacc acattatccc aactggttta aagccctgag gatagtagtc actttattgg 840
ccactttttc atttataatc acttttagtag ttaaaagtag ttttccagaa aaaggacata 900
agagacctgg tcaagtataa aaaatataaa agtctgggaa gtgaggagca cctctgcccc 960
gctgctgccc cgtctgggaa gtgaggagcg cctctgcctg gccgcctgac catctgggaa 1020
gtgtgacaag cgcctctgcc cggccgctgt gcaaccttcc acgtgtgaag tgacagcctt 1080
gtgtgtgatc ttttctgtct tcccgaagt tgcatcttgc acattaaagt ttacttttta 1140
gttaaaagtt ttaaaaatat atatatataa atacactgta gataacattt gtatgccagc 1200
tacacctttt tctacttctg tttggctttt tttcccaca ccaatggtaa tttatcttca 1260
cagattgttc ttcatcttca gaaattgtta cttcatgtta attacttgag caaaagcttg 1320
aaaatccctg acaagtactt ttcatctcat agtatattag ttttcaactca gtcattttat 1380
gaataatata gttatccact taaacatttc aatattttta ccatcttgaa aattaaagat 1440
taaaaatccc aaaaaaaaaa aaaaa 1465

```

<210> 56  
 <211> 1170  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3145660CB1

```

<400> 56
ctctcataca gaagctagct gcttgttttt tcactttatc tattatcgga ttttaattt 60
tttacataat tcactgtaat gctcatcgga agaataagcc gtgtactgat ttggaaagt 120
gagaggaaaa gaaaaatatt atcaataccc ctgtgtcttc attagaaata cttttacaat 180

```

PF-0681 PCT

```

ctttctgcaa acttggcctg attatggcat atttctatat gtgtgaccgt gcaaactctgt 240
tcatgaagga aaacaaattt tatacacatt catctttctt tattccaatt atctacattt 300
tgggttttggg agtattttat aatgaaaata ctaaagagac taaagtatta aatagagaac 360
aacagacga atggaaaggc tggatgcaac ttgtgatttt gatttatcac atttctggag 420
caagtacatt tttgcctgta tacatgcaca ttcgagttct gggtgctgca tattttattt 480
agacagggtg tgggcatttc tcatactttt ggataaaagg agattttgga atctatagag 540
tatgtcaggt tttatttcgt ctcaatttcc tggtagtggt gttatgtata gtaatggatc 600
gaccttatga attctattac tttgtcccct tggtcactgt atgggttcag gtcatatatg 660
ttacttttagc actatggcca caaataatcc aaaaaaaggc aaacggaaat tgtttctggc 720
attttggtct actgttgaaa ctaggctttt tgctgttatt catatgtttt ttggcatatt 780
ctcagggtgc atttgagaag atcttttctc tttggccatt gtccaagtgt tttgaactga 840
aagggaatgt atatgaatgg tggttcagat ggagggttaga ccgttatgta gttttccacg 900
gaatgctgtt tgcttttatt tatctggctt tgcagaagcg tcaaataact tctgaaggaa 960
aggggtgaacc tcttttttca aacaaaattt caaattttct gttggttatt tcagtagttt 1020
ctttcctggg taagttttga aaacctttcc caaattctaa aatgggtatg ggaatattta 1080
accattttaag ataatggaaa atataaatcc aattattttt taataattaa atcctttttt 1140
tcctccctta tttttccctt ttaataattt

```

&lt;210&gt; 57

&lt;211&gt; 1421

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 4901066CB1

&lt;400&gt; 57

```

gaaccaggaa agtaacggct acagacagtg agaaatagtt tcgctcgccg gctagaaaaa 60
ctctgtcggg accaacccca gagcgttgag agcagccccc ctccacgctt ccttaacgga 120
gaggtgcagg actcagactt caccagccca ctcggtccca gccttgtagc caaagagacg 180
ccaaggacgc gctctcccgc gtccaggcag cccagcttg ctggcttgcc tgcccgcctg 240
cgtgcagcac tcggccggcg tgcagcatga cctgtggaa cggcgtagtg cctttttacc 300
cccagccccc gcatgccgca ggcttcagcg ttccactgct catcgttatt ctagtgtttt 360
tggctctagc agcaagcttc ctgctcatct tgccggggat ccgtggccac tcgctgtggt 420
tttggttggt gagagtctt ctcatgtctg tcataggcgc agaaattgtg gctgtgcact 480
tcagtgcaga atgggttcgt ggtacagtga acaccaacac atcctacaaa gccttcagcg 540
cagcgcgcgt tacagcccgt gtcggtctgc tcgtgggcct ggagggcatt aatattacac 600
tcacagggac cccagtgcag cagctgaacg agaccattga ctacaacgag cagttcacct 660
ggcgtctgaa agagaattac gccgcggagt acgcgaacgc actggagaag gggctgccgg 720
accagtgcgt ctacctggcg gagaagttca caccgagtag cccttgccgc ctgtaccacc 780
agtaccacct ggccgggacac tacgcctcgg ccacgctatg ggtggcgctt tgcttctggc 840
tcctctccaa cgtgctgctc tccagcccg cccgcctcta cggaggcctg gactgctga 900
ccaccggagc cttcgcgctc ttccgggtct tcgcttggtc ctccatctct agcgtgccgc 960
tctgcccgct ccgcctaggc tcctccgcgc tcaccactca gtacggcgcc gccttctggg 1020
tcacgtggc aaccggcgct ctgtgcctct tcctcgagg ggccgtggtg agtctccagt 1080
atgttcggcc cagcgtctt cgcaccttc tggacaaaag cgccaaggac tgagccagg 1140
agagaggggg ctacactctt atcctcggcg acccactgca caagcaggcc gctctcccag 1200
acttaaaatg tatcaccact aacctgtgag ggggacccaa tctggactcc ttccccgct 1260
tgggacatcg caggccggga agcagtgcgc gccaggcctg ggccaggaga gctccaggaa 1320
gggcactgag cgctgctggc gcgaggcctc ggacatccgc aggcaccagg gaaagtctcc 1380
tggggcgatc tgtaataaaa ctttttttt ttttgtttt t

```

&lt;210&gt; 58

&lt;211&gt; 2310

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 5031174CB1

PF-0681 PCT

&lt;400&gt; 58

```

ccggcccttag gtccccggcag cgggtggtgac ggcgggtgccg gaggttgtcc ttggcaggtt 60
ttctctcgccg cttctccatg gaggaggcgg tgcgaacggc ttcagccccg aatgctcgca 120
tctcccactg gacggcgacg aaggcgggtg ccgtgcgagc gcaggactgg gcggcctgtg 180
tggggggtgtg agccgcggtg cccaaggctg cgcggcgag ggggaagccgc gcggccggcc 240
ggccgactag ggtttgaaga ctgcttcatt ctgcctctag taccagcggg ttctctgttc 300
tgtgatcaat gtgattcaca ggaactcctt aagtaacaaa cgaaatgagc caggggcgtg 360
gaaaatatga cttctatatt ggtctgggat tggctatgag ctccagcatt ttcattggag 420
gaagtttcat tttgaaaaaa aagggcctcc ttcgacttgc caggaaaggc tctatgagag 480
caggtcaagg tggccatgca tatcttaagg aatggttgtg gtgggctgga ctgctgtcaa 540
tgaggctgg tgaggtgccc aacttcgctg cgtatgcgtt tgcaccagcc actctagtga 600
ctccactagg agctctcagc gtgctagtaa gtgccattct ttcttcatac tttctcaatg 660
aaagacttaa tcttcatggg aaaattgggt gtttgctaag tattctagga tctacagtta 720
tggtcattca tgctccaaag gaagaggaga ttgagacttt aaatgaaatg tctcacaagc 780
taggtgatcc aggtttttgtg gtcttttgcaa cccttggtgt cattgtggcc ttgatattaa 840
tcttcgtggg gggctcctcgc catggacaga caaacattct tgtgtacata acaatctgct 900
ctgtaatcgg cgcgttttca gtctcctgtg tgaagggcct gggcattgct atcaaggagc 960
tggttgagg gaagcctgtg ctgcggcctc ccctggcttg gattctgctg ctgagcctca 1020
tcgtctgtgt gagcacacag attaattacc taaatagggc cctggatata ttcaacactt 1080
ccattgtgac tccaatata t atgtattct ttacaacatc agttttaact tgttcagcta 1140
ttctttttaa ggagtggcaa gatatgcctg ttgacgatgt cattggtact ttgagtggct 1200
tctttacaat cattgtgggg atattcttgt tgcattgcct taaagacgtc agcttttagt 1260
tagcaagtct gcctgtgtct tttcgaaaag acgagaaaagc aatgaatggc aatctctcta 1320
atatgtatga agttcttaat aataatgaag aaagcttaac ctgtggaatc gaacaacaca 1380
ctggtgaaaa tgtctcccga agaaatggaa atctgacagc tttttaagaa aggtgtaatt 1440
aaaggttaat ctgtgattgt tatgaagtga atttgaatat catcagaatg tgtctgaaaa 1500
aacattgtcc tcaaataatg ttcttttaaag gcaatctttt taaagatttc actaatttgg 1560
accaagaaat tacttttctt gtattttaaac aaacaatggt agctcactaa aatgacctca 1620
gcacatgacg atttctatta acattttatt gttgtagaag tattttacat tttcatccct 1680
tctccaaaag ccgaatgcac taatgacagt ttttaagtcta tgaaaatgct ttattttttc 1740
attggtgatg aaagtctgaa atgtgcattt gtcacccca ctccatcaat ccctgaccat 1800
gtaaggcttt tttattttta aaaaacagag ttatcccaat acattatcct gtgatttacc 1860
ttacctacaa aagtggctcc tgtttgtttg atgatgattg gttttatttt tgaaatattt 1920
attaagggaa aactaagtta ctgaatgaag gaacctcttt cttacaaaac aaaaaaagg 1980
gcagaaatca cccaaggaa cgattttctca ggttgagatg atcaccgtga atccggcttc 2040
ctctgagcat tcgatggcct tagcacctca tcaagccagc acatcctgcc tgctgttgca 2100
gcctggctgg gtttattctt cagttaccct aatcccata tgccctggaac cttgattacc 2160
gttttacatc agctcttgta cttttcagta tattttcata atgagttata ttgtcattta 2220
gactttgaac agctctggga aatagaagac taggggtgtt tcttaaattt agctcatgtt 2280
ataataaaaa gttgcaaaat aaaaaaaaaa 2310

```